

The end of the American Dream?

The stock-market crash of the past two weeks signals the end of more than a mere bull-market. Even if governments have the wit to avoid a serious depression, United States citizens will be impoverished.

THE crash on the world's stock markets in the past two weeks seems to have brought out the worst in people who should have known better. There seems to have been a competition to emulate President Theodore Roosevelt's fatuities on the similar events of 1929: President Ronald Reagan's description, two weeks ago, of the slump in stock prices as a "readjustment" has now been echoed by the president of the British Confederation of British Industry. Stock-market managers have also been blaming their machinery for their misfortunes. Thus the New York Stock Exchange has been pleading with its traders not to let their computers generate orders to sell stocks, but to do the job manually instead. There has also been an outburst of indignation at the use of the world's markets in the future prices of stocks as a means of insuring against rapid changes in the value of publicly traded stocks. Are the stock markets really saying that there would have been no crash if their customers had been less able to act quickly, and denied the hope of protecting themselves against loss?

The temptation should be resisted to pretend that this wild talk, and the events to which it refers, concern only the inward-looking community of those who deal in other people's money. The events of the past two weeks concern us all. First, there is the near-certainty that most of us will turn out to have lost money; the pension funds and insurance companies which have grown to be the big players on the world's stock markets represent the interests of broad sections of the world's free-market communities. Their losses in the past two weeks are everybody's losses. Moreover, there is no great substance in the argument that the losses are but paper losses, and that, because the financial institutions still between them own the same proportion of industrial assets as before the crash, nothing has been permanently lost. For one thing, those lucky enough to have turned their share of paper wealth into cash before the crash (perhaps because they had quit working) have benefited at the expense of others. More important, the prices being paid for stocks (and thus, indirectly, for the productive assets they represent) will quickly be discovered to have been uneconomically high. The problem is especially serious in the United States, where even a successful navigation of the troubles ahead will require important and painful shifts in the pattern of the US economy.

Recession

The second danger is that matters will be made worse than they are already by a serious economic recession, perhaps even by a slump along the lines of the 1930s. There is nothing remarkable in this observation; indeed, the way in which central banks in Western Europe and the United States have been reducing interest rates, while a necessary response to the plight of the stock markets, shows that the authorities are also aware of the more distant danger. But there are limits to what governments can do. After the crash, people will think themselves less wealthy than they were before (which will be literally true for those wishing to turn their stock-market assets into cash or other kinds of assets, such as dwellings), and so will spend less, reducing the demand for goods and services. Only time will tell how important this effect will be, but the assumption so far that rates

of economic growth will be reduced by a third or thereabouts could be wildly optimistic. Events may yet show that, left to themselves, growth rates could become negative.

Free-market governments all know what should be done when there is a prospect of recession. This is when Keynes comes into his own. Governments should spend more, or alternatively reduce taxes, so that the demand for goods and services is increased, and should be unafraid of temporarily running their own budgets at a deficit. The embarrassment, in the United States, is that the government took precisely those steps seven years ahead of time — and for the past three years has been trying vainly to reduce the deficit to more manageable proportions. Only last week, President Reagan offered the US Congress negotiations on a strategy eventually to win a balanced budget. In logic, he could have said that this is not when the US government should be drawing in its horns. But the accumulated deficits of the past several years, and the present fear that the US government has lost control of fiscal affairs, leave no room for manoeuvre. If the deficit were allowed actually to grow, it could not be financed as in the past few years by overseas investment in the United States, but only by printing the missing money. The United States is boxed in between inflation and recession.

The way out must be painful. The counterpart of the domestic budget deficit is the adverse external trade balance, which in August exceeded \$15,000 million, roughly the equivalent of \$60 for every US citizen in that month alone. The chances are that, with the precipitous fall of the dollar against most other currencies in the past ten days, the trade balance will have been corrected, which will mean that US citizens will feel (and be) \$60 a month worse off. The problem for the future is that of paying off the accumulated debts without the pursuit of protectionist policies (which could also precipitate a depression). That means that the trade balance must be turned round, requiring first enough savings from people's income to sustain fresh investment in manufacturing plant and then a steady stream of exports. The result will be a transfer of personal prosperity from the United States to its trading partners. That dismal prospect is the best possible outcome from the present box.

Do governments have the will and the means to plot such a course? The record is not encouraging. Most recently, sensing that the imbalances were getting out of hand, governments and central banks joined forces to prevent what has since happened — a further substantial fall in the value of the US dollar. That should have been a warning to those concerned that events had escaped from their control. It is worrying that the warning seemed ineffective, and does not promise well for the difficult months ahead, when all free-market governments will have to strike a judicious balance between firm and free economic policy. With a presidential election in the offing, the US government will need particularly strong nerves to superintend what is bound to seem the impoverishment of the United States. The task will be especially difficult because even the wisest and most courageous policies may be defeated by bad luck — the failure of a major financial institution or the bankruptcy of a major debtor among the developing countries. Yet the alternative is worse. □



Japanese industry rallies around superconductors

- New research centre planned
- International cooperation hopeful

Tokyo

JAPAN'S Ministry of International Trade and Industry (MITI) looks set to succeed in persuading the giants of Japanese industry to join forces in researching the recently discovered high-temperature superconductors. A massive new research centre, funded by industry, will be established and will help to provide a total national investment in superconductor research that will at least equal that of the United States.

First news of the centre was leaked in the summer from closed sessions of MITI's high-temperature superconductor research committee, a key advisory group whose members are drawn from industry and academic institutions (see *Nature* **328**, 282; 1987). According to Professor Shoji Tanaka of Tokyo University, one of the committee members, the centre will be divided into an information centre and a research laboratory.

Eighty to ninety companies are interested in joining the information centre, says Tanaka, and each will contribute an annual subscription of \$10,000 to give the centre a million-dollar-a-year budget. Along with the electric utility, electronic and cable manufacturing companies, securities companies, banks and construction companies are among potential members. The latter are looking for investment opportunities and for the big construction works that would be necessary for magnetically levitated trains.

But the really big money will go to the centre's research laboratory. Some 35–38 companies, including such electronics giants as Toshiba and Hitachi, will participate, according to Tanaka. Each will contribute about ¥100 million (\$700,000) at the start and ¥20 million (\$140,000) a year for running costs.

Setting that money alongside the approximately ¥6,000 million (\$42 million) in superconductor research spending that the two big technology agencies, MITI and the Science and Technology Agency, are asking for in their 1988 budgets puts Japan at least level with the United States in its total commitment to superconductor research. US government and industry will together spend \$117 million in 1988, according to Charles Laverick, a consultant for Westinghouse Research and Development Center and one of the authors of the US National Academy of Sciences report on superconductor

research. And many believe that estimate may be optimistic.

The new centre will apply to MITI for classification as a non-profit organization, so avoiding corporate and property taxes, according to Rizuhiro Nezu of MITI's Agency of Industrial Science and Technology. He says the information centre is expected to be in place in Tokyo within the next two or three months, a little later than had been thought, while the laboratory should be built by next summer.

How will the new centre be viewed by Japan's foreign competitors? It will no doubt stir painful memories of the very-large-scale integrated circuit research project which, in 1976, brought together competing electronics companies and government laboratories under MITI's sponsorship. Within a year of the project's completion in 1980, Japanese companies had seized 70 per cent of the world market for memory chips, previously dominated by US companies. But this time there is a big difference, says Tanaka. The laboratory will be concerned primarily with basic research and, if Tanaka and Nezu have their way, it will be open to participation from foreign companies and researchers. However, Nezu admits that no attempt to approach foreign companies has yet been made.

Tanaka, best known for his group's confirmation of Bednorz and Müller's discovery of high-temperature superconductivity, is also one of Japan's leading semiconductor researchers. And he is most concerned that superconductor research should not follow semiconductor research as a source of a bitter trade dispute with the United States. But, although Tanaka strongly favours international co-operation, he says Japan must be "very careful" about the military aspects of some US research.

Much will depend on attitudes in the United States too. The US administration sees it as vital that the United States should win the superconductor race in order to restore confidence in the nation's ability to innovate. And there is an entrenched belief that joining forces with Japan would only make it easier to export US basic research. Others, however, estimate that practical applications of the new superconductors are decades away — except in a few specialized areas — and see a golden opportunity for long-term international collaboration in truly basic research.

David Swinbanks

Antarctic not the place for sun worshippers

Washington

ANTARCTIC explorers are going to have to pack plenty of sun-tan oil along with their dog-sleds if the latest calculations of the effects of the depletion of the ozone layer prove correct.

According to Peter Wilkniss, director of polar programmes at the US National Science Foundation, a day under the Antarctic ozone hole will provide a dose of ultraviolet radiation four times as great as a day on the beach at Miami.

Wilkniss, who gave evidence to a Senate Environment and Public Works Committee last week, says he finds himself in the difficult position of being "responsible for the first US population living under a depleted ozone shield" and just does not know what measures are going to prove necessary. All he has to go on are some preliminary calculations that assume a worst case in which stratospheric ozone levels fall to 100 dobson units, the expedition crew are outside for the polar day and the Sun is high in the sky.

The worst-case conditions may not be overly pessimistic: satellite observations show ozone levels of 105 dobson units over the US McMurdo station (although they have yet to be squared with simultaneous ground records of 127 dobson units), and the Sun does rise well above the horizon at the more northerly stations, although not at the South Pole.

Wilkniss wants to obtain data on Antarctic ultraviolet levels as quickly as possible. Measurements will soon be available from equipment which was installed at the South Pole last year, but they are not likely to provide a complete answer: it is still dark at the South Pole when the ozone hole is at its deepest. Next year, monitoring programmes will be under way at all US stations and a safety panel will assess what precautions are necessary.

Part of the problem in estimating the dangers, Wilkniss says, is that crew who over-winter in the Antarctic spend 6–8 months without sunlight. That is known to affect their hormonal and immune systems and may render them more susceptible to harm from high ultraviolet light levels when the polar spring arrives and the ozone hole deepens.

The biggest immediate danger is to the eyes. Temporary 'snow blindness' is already quite common and extra protection against increased ultraviolet light may be necessary. Further in the future are heightened risks of skin cancer.

Alun Anderson

United opposition to three-tier science in UK higher education

London

PROPOSALS to create a three-tier system of science research and teaching within British higher education have brought almost universal condemnation from the scientific community, with representative bodies united in their opposition.

Last July, in response to a request from the government, the Advisory Board for the Research Councils (ABRC), whose job is to advise the Secretary of State for Education and Science on his responsibilities for civil science, produced *A Strategy for the Science Base*, advocating far-reaching changes in the organization, management and support of scientific research in higher education and the research councils (see *Nature* 328, 280; 1987).

The most controversial of the proposals were that institutions be categorized into three types, with some being recognized as primarily teaching institutions, with no advanced research facilities; that funds be transferred from the University Grants Committee to the research councils to contribute towards the costs of research overheads; and that a series of interdisciplinary university research centres be established, linked with academic institutions not in the 'teaching category'. The government invited comments on the proposals from 'interested bodies'.

The concept of institutions devoted solely to teaching has provoked by far the most vehement opposition. The Royal Society says that despite ABRC's contention that a three-tier system could be sufficiently flexible to allow institutions to change their status, a formal definition of institutions would be final.

The Committee of Vice-Chancellors and Principals (CVCP) agrees that "no university should be restricted to a teaching role alone". The Association of University Teachers (AUT) describes the proposal as "absurd" and says that the proposed categorization is "incompatible with the objectives of university education".

While there is a broad consensus that greater concentration of research is needed, there is disagreement about the best means of achieving it. The CVCP welcomes the concept of interdisciplinary research centres as having "potentially important parts to play in the concentration process", but is concerned to know details of how the centres would be managed and funded. The Royal Society recommends that no more than 10 per cent of research council funds is channelled through any such centres. The AUT is suspicious of the proposals, particularly on the question of the management of the centres, fearing that they "may obtain autonomy, not subject to the

shared academic objectives for the institution set up by its senate, and in the longer term may set themselves on the road to hiving-off into autonomous private institutions".

The AUT would prefer to see 'interdisciplinary research groups' develop through collaborative short-term project grants, leading to longer-term programme grants over a period of, say, five years "rather than being arbitrarily planted".

A joint statement from a group of scientific societies and institutions rejects the concept of interdisciplinary research centres. "This degree of selectivity has often failed in the past and there are few reasons for believing that it would be more successful in the future." The British Association for the Advancement of Science welcomes the concept of such centres as places to encourage voluntary cooperation between scientists within single institutions or between institutions.

Transfer of funds from the University Grants Committee to the research councils to pay for research overheads does not find much favour, with the AUT claiming that such a move would challenge the autonomy of the universities and would result in less effective targeting of funds.

Last week, the Secretary of State for Education, Mr Kenneth Baker, made clear his commitment to reform in the university system when he met vice-chancellors.

Baker said that he did not believe the "old order" of higher education could have lasted. "Stumbling on in the old ways would in the long run do far more harm than radical change." He gave cause for tentative optimism, however, when he told the vice-chancellors that "the success of our policies is dependent on your cooperation."

The government's response to the ABRC document is expected early next year. It seems unlikely that any of the recommendations will be incorporated into the forthcoming education bill, whose publication is expected within the next few weeks.

Simon Hadlington

United States cheered by its successful Titan launch

Washington

LAST week's successful launch, after two failures, of a Titan rocket from Vandenberg Air Force Base in California does not amount to a resumption of the US space programme, but coming after an equally

the type designated KH-11. The reconnaissance capabilities of the US Department of Defense have suffered enormously from the Titan and shuttle failures, and there is speculation that the launch, which was unannounced, succeeded in replacing the last active KH-11 only weeks before its anticipated 3-year lifetime expired.

In the civilian space programme, the next four scientific missions planned by the National Aeronautics and Space Administration (NASA), the Magellan, Galileo, Ulysses and Mars Observer projects, are all intended to use the shuttle. (NASA announced on 22 October a revised shuttle launch schedule, reducing the number of flights in the first year, but the slots for the planned science missions have been maintained.) The earliest use expected of a Titan rocket is not until the spring of 1993, for an as-yet unapproved comet rendezvous, although there is still a possibility that Mars Observer could be sent aloft on a Titan. For Geoffrey Briggs, NASA's director of Solar System exploration, the wider significance of last week's launch lies in the availability of the Titan as a back-up launch vehicle should the resumption of shuttle flights suffer further delays. Even before the Challenger disaster, there was criticism of NASA's reliance on a single launch vehicle, and although Briggs says that some problems still remain, NASA has learned its lesson and is actively seeking to maintain a varied launching fleet.

David Lindley



successful test of the redesigned space shuttle motors (*Nature* 329, 93; 1987), it does at least engender new confidence in the ability of the United States to put things into orbit. And the effort that has gone into the development of the Titan as a reliable launcher is a sign that both the civilian and military space programmes have recognized the danger of depending solely on the space shuttle.

The military payload sent into orbit on 26 October was not identified by the Air Force, but was reportedly a spy satellite of

UK's long-held scepticism of European space plans

London

BRITAIN'S scepticism of the proposed long-term programme of the European Space Agency (ESA) to be debated by ministers at The Hague next week is not confined to members of the government. Officials at the British National Space Centre (BNSC) have also become concerned at the escalating costs of the programme and the questionable merit of heavy investment in manned projects.

It has been widely assumed that it was



Gibson — is ESA out of its depth?

solely government parsimony that prompted Britain's objections. But last week Mr Roy Gibson, the former director-general of BNSC who resigned in August after the government refused to increase the space budget, told the House of Lords Select Committee on Science and Technology that during discussions with ESA officials over the past 18 months, UK delegates have consistently warned that the proposed programme is too ambitious.

According to Gibson, the rapid evolution of ESA's programme takes it well beyond the bounds set by ministers in Rome in 1985. Gibson feels that ESA is biting off more than it can reasonably chew. "We have had the feeling that ESA has not realized the extent of the commitment it was taking on."

One of the main concerns expressed by UK delegates is the growing imbalance in the programme in favour of large infrastructure projects at the expense of user programmes. Gibson questions the accuracy of some of ESA's cost-of-completion studies, in particular that of the Hermes space vehicle. Initially, ESA quoted £2,000 million; Gibson, who describes Hermes as "a luxury", argued that the true cost would be nearer £7,000 million. ESA has since revised its estimate to £4,000 million. Gibson feels that complete package ESA's will not be adopted.

Nobel error

IN "MIT needs millions" (*Nature* 329, 757; 1987) the 1987 Nobel prize for economics was mistakenly awarded to David Saxon, MIT's chairman.

The real recipient is, of course, Professor Robert M. Solow of the same institute.

but that a seriously imbalanced programme could nevertheless emerge.

It is ironic that Gibson's view of the ESA proposals coincides with that expressed by the government, which still refuses to increase the £116 million Britain now spends annually on civil space research.

Simon Hadlington

● British ministers are generally badly equipped to determine the guidelines of space policy, the systems used by Whitehall for garnering technical advice are imperfect and most officials have little grounding in the scientific and technical substance, claims a study by the Royal Institute of International Affairs, published this week*. "Space policy has been only one of many examples of drift and confusion in policy-making for science and

technology, compounded by the sheer difficulty of comprehending what was at stake."

The report says the country's fragmented approach to space policy and its failure to demonstrate commitment to international collaborations has prevented it from exploiting technological and political advantages.

The report warns that if Britain does not increase its contribution to ESA it will not be able to retain the national capabilities it would need if security interests and technological developments began to favour a larger military programme, either nationally or within a European context. "Britain cannot expect to retain an effective capability in space at the current level of engagement . . . Space has far-reaching implications for British economic modernization as well as for Britain's foreign and security policies." □

*British Space Policy and International Collaboration. (Royal Institute of International Affairs, London. £6.95.)

Protests in South Africa over threatened subsidy cuts

Cape Town

THOUSANDS of students, staff and alumni took part in general assemblies at South African universities on 28 October to protest at the imposition of the subsidy conditions promulgated on 9 October by the Minister of National Education, Mr F.W. de Klerk. Assemblies were held at the Universities of the Witwatersrand, Cape Town, Western Cape, Natal and Westville, and at Rhodes University the previous week.

Academic processions round the perimeter of each campus were held without incident, except at the University of the Witwatersrand where riot police baton-charged students at the edge of the campus in Johannesburg's Jorissen Street. At the Universities of Cape Town and the Western Cape, exits from the campus were manned by police and at Western Cape police helicopters circled over the 5,000-strong procession.

These protests form part of a joint response to the subsidy conditions. In August, the universities were asked to respond to a set of government proposals requiring university councils to prevent various kinds of political activity on campuses, including specifically interference with academic activities, unlawful gatherings, promotion of strikes, boycotts or work stay-aways, actions of civil disobedience or the printing of any publication in contravention of the internal Security Act. These proposals were rejected by several of the country's universities (see *Nature* 329, 192; 1987). This opposition has been largely ignored by de Klerk, as the new regulations are virtually the same

as his original proposals, with the exception of a provision curbing "affected" organizations, which has been omitted. The universities are said to be considering appealing to the courts to declare the regulations *ultra vires*. The regulations themselves are not enshrined in an act of parliament, and have been promulgated by the minister in terms of the Universities Act, 1955.

The rector of Pretoria University, Professor Danie Joubert, said that he was not particularly concerned about the regulations, as he was happy that the government intended to accede to his request that it "acknowledge the authority of the university council". He had received assurances that in the event of a contravention of the regulations, the government would not unilaterally impose a subsidy cut, but would first enter into consultation with the university concerned.

The president of the Council for Scientific and Industrial Research, Dr Chris Garbers, said that he did not feel that he was in a position to comment on the regulations, as he had not been a party to the negotiations between the government and the universities and therefore did not have first-hand knowledge of all the relevant details.

Meanwhile, at Stellenbosch University, about 200 students marched to the office of the rector, Professor Mike de Fries, to deliver a protest deploring his acceptance of the regulations, which it described as "an attempt by the government to turn the universities into instruments to sustain the National Party idea of law and order".

Michael Cherry

Soviets admit improper committal to mental hospitals

London

THE Minister of Health of the Russian Soviet Federated Socialist Republic, A. Potapov, has admitted that some Soviet citizens have been improperly committed to mental hospitals on account of their outspokenness. In the past, the Soviet psychiatric establishment has consistently denied allegations that psychiatric methods were being applied against political dissidents.

This is still the official stance — and although it was recently reported that a

law is being prepared making the political misuse of psychiatry a crime in the Soviet Union, some Soviet commentators have paradoxically presented the promised law

Now, don't you feel better for admitting that?



as a proof that such abuses do not exist. An exposure, earlier this year, by the newspaper *Sotsialisticheskaya Industriya*, of improper conduct by some psychiatrists dealt not with political cases but with bribe-taking by certain psychiatrists, who certified criminals as mentally incompetent so that they could avoid long prison sentences, or providing young men with certificates to avoid military service.

Sotsialisticheskaya Industriya subsequently received details of cases of the hospitalization of people whose only symptom was to have been overly vocal in their criticism of poor working conditions or economic abuses. In a lengthy interview with the paper, Potapov, himself a psychiatrist, was asked to comment on two such "revenge" committals, in one of which the "patient" was hospitalized on the basis of a work-performance report and two statements by people whom the doctor did not even meet, while in the other, a psychiatric ambulance was summoned to the work-place to collect someone who had expressed some apparently well-founded criticisms.

Potapov replied that the existing regulations on committals should make such cases impossible, "but sometimes the instructions are carried out by people who, regrettably, do not always come up to the necessary professional standard". Several cases of "arbitrary" applications of the rules, leading to "violations of human rights and the sustaining of moral traumas" had been substantiated, he said, and it would therefore be "expedient" to revise the rules so as to make such "broad interpretations" impossible.

In a Soviet context, such an admission is extremely significant. In 1983, the Soviet All-Union Society of Psychiatrists and

Neuropathologists withdrew from the World Psychiatric Association, just in time to forestall its expulsion on grounds of psychiatric abuse. Since then, the society has put out discreet feelers about the possibility of readmission, but has been given to understand, equally discreetly, that colleagues abroad would first want to see good evidence that psychiatric methods of repression were no longer in use against dissidents. The standard Soviet rejoinder is that claims of psychiatric abuse in the Soviet Union are unsubstantiated, and are simply a ploy of anti-Soviet campaigners abroad.

Potapov's more cautious approach, admitting that there have been certain infringements, including cases of critics of abuses (which is how most "dissidents" began their careers) seems more liable to evoke credibility abroad and be a more effective road to eventual reinstatement.

Vera Rich

Celltech undeterred by crash of '87

London

THE Slough-based biotechnology company Celltech is going ahead with plans for the biggest-ever private international share placing, despite the massive falls in the world's stock markets during the past three weeks. As a private company, Celltech is not directly affected by market fluctuations, and the company is confident that the £50-million placing will succeed.

Celltech was set up in 1980 with capital of £12 million from City institutions and the government. The company has grown steadily since — its approximate value was £94 million in August, before the 're-adjustment' in company valuations — but as with other biotechnology companies, the profits are yet to come. In the year to September 1986, sales were £11 million for a loss of £714,000, but the company expects to make a profit this year.

Most of Celltech's income has come from monoclonal antibody production for diagnostic applications and, until now, products have been licensed at an early stage to major pharmaceutical companies for product development and clinical trials. The hoped-for £50 million is intended to enable the company to maintain European marketing rights in the most promising products resulting from in-house research, and to go on to the development and clinical-trial phase on its own behalf.

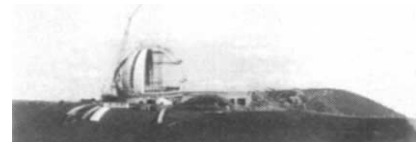
Perhaps to help the new share placing along, Celltech chose last week to announce a new contract from Ortho Pharmaceuticals which should bring in several million pounds over the next two years. Ortho recently won a licence from the US Food and Drug Administration to market the hormone erythropoietin for treating renal failure and to investigate its applicability in other diseases. Celltech's role will be to assist in the production of erythropoietin for the clinical trials and early marketing.

Charles Wenz

Keck telescope

San Francisco

CONSTRUCTION of the dome for the world's largest infrared telescope, the 10-m Keck Telescope, on Mauna Kea in Hawaii, will be finished before the winter snowfall. The \$87-million telescope is being built as a joint project of the Uni-



Dome completion soon.

versity of California and the California Institute of Technology.

The telescope will be installed in the autumn of 1988, and assembly and alignment of the telescope's unique honeycomb array of 36 hexagonal mirrors will begin in 1989. The mirror arrangement was conceived as a way to circumvent the distortion and sagging that would necessarily arise because of the weight of a single 10-m mirror.

Two new technologies have been developed for the telescope, a computerized monitoring and adjustment system to keep the mirrors perfectly aligned, and a 'stress mirror polishing' procedure to produce the asymmetrical surface required of each mirror in order to give the final, symmetrical composite surface.

Project spokesman John Gustafson says two mirror segments have been polished so far, but polishing has been halted until final calibration is completed on the device that will test the mirrors for the proper shape. Despite the calibration delays, Gustafson says the project is still on schedule, with plans for the telescope to begin operation in early 1991.

Marcia Barinaga

NOAA at sea: who will be named to take over the helm?

Washington

HAVING lost its administrator and with no replacement yet named, it is again being asked on Capitol Hill whether the National Oceanic and Atmospheric Administration (NOAA) can function effectively as part of the Department of Commerce.

Although it receives about half the department's budget, its research and weather forecasting functions have always seemed out of place in an agency that is also responsible for economic development and trade policy. But despite interest from Congress, NOAA is unlikely to be moved out of Commerce, and its new administrator will have to continue to chart its course through treacherous bureaucratic waters.

NOAA is not a small agency. Its 1987 operating budget was approximately \$1,160 million, not far behind the National Science Foundation's \$1,620 million budget. It is the parent agency for the National Weather Service, the National Marine Fisheries Service and the National Ocean Service. It also sponsors satellite data collection activities and supports research programmes on the ocean and atmosphere. But the Reagan administration has not been a strong supporter of NOAA's scientific activities, consistently trying to reduce or eliminate NOAA research programmes.

Commerce secretaries — by organizational charts the overseers of NOAA — have tended to be chosen more for their knowledge of trade issues than of NOAA activities. Indeed, in hearings held before his confirmation by the Senate last month, Secretary C. William Verity confessed that before he was asked to head the department he was not sure what NOAA did.

Part of NOAA's financial woes stem from the way its budgets are analysed within the White House. Despite its scientific activities, NOAA's budget is not reviewed at the Office of Management and Budget by the staff that deals with other science agencies. NOAA administrators have a chronic need to educate those who will pay for their programmes.

But NOAA does have its allies. In Congress, Senators Lowell Weicker (Republican, Connecticut) and Ernest Hollings (Democrat, South Carolina) are supporters of the agency. Weicker believes that NOAA should be given the status of an independent agency, like NASA (National Aeronautics and Space Administration) or EPA (Environmental Protection Agency). Although he has introduced legislation that would require that move, the bill is unlikely to go anywhere during this session of Congress. Support for making NOAA independent is not univer-

sal: some fear that an independent agency would be more vulnerable to budget cutting, as, unlike NASA, it does not have a strong public identity.

The White House is expected soon to announce its choice for a new administrator to replace Anthony J. Calio who left NOAA last August. Calio resigned as administrator just one week after Malcolm Baldrige's death in the summer. He was widely regarded as having been an effective force for the agency within the administration. Calio says he developed a close working relationship with Baldrige, and was personally shaken by his death in a rodeo accident.

Two names being mentioned as possible successors to Calio are Carl Savit, recently retired from Western Geophysical in Houston, and William E. Evans, currently assistant administrator for fisheries at NOAA.

Joseph Palca

Czech demand for modern computers

London

CZECHOSLOVAKIA'S import regulations governing personal computers are ill-conceived and encourage technical backwardness, according to the Czech party daily, *Rude Pravo*. The present regulations base the duty payable on the memory capacity, which, the paper admits, may seem "simple and reasonable", but which is questionable on two counts.

First, it is not the memory capacity that determines the capabilities of a personal computer, but the speed of operation. Second, the rate of duty is set at 150 Czechoslovak crowns per kilobyte, which becomes prohibitive for the new megabyte memory personal computers.

The demand for personal computers in Czechoslovakia far exceeds the supply. Domestically produced ones are manufactured in very short production runs, the paper notes, and are based on outdated components (for example, the 8080 integrated circuit, or the Z-80 which is only slightly more advanced).

Personal computers and minicomputers are particularly needed, the commentator points out, in the educational sector, from primary schools to universities. Even the specialized primary school for the mathematically gifted has only 14 computers for the entire school. The situation is worst, *Rude Pravo* notes, in technical universities, "although it is precisely the future engineers who should be working with the latest computer technology while they are studying, rather than being introduced to it later".

Vera Rich

US agents claim success over supercomputers

Washington

US OFFICIALS say they have plugged a potentially disastrous technology leak by arresting three men accused of trying to sell the Soviet Union plans of a supercomputer being built by Saxpy Computer Corp of Sunnyvale, California. Defence officials say the stolen technology has "a direct military application and would have proven extremely detrimental to US national security if it had gotten to the Soviets".

Sandy Towles, director of market development for Saxpy, said the MATRIX 1/1000 computer, which sells for \$1.9 million, would have been attractive to the Soviet Union because it is constructed primarily from readily available hardware, much of which would not be subject to export control.

The supercomputer is designed for numerically intensive computational environments. It uses SIMD (single instruction, multiple data) architecture, and can execute as many as 1,000 megaflops (million floating point operations) per second for certain applications. Saxpy has installed one computer at Martin Marietta Corporation in Baltimore, and has sold a second machine to the Canadian company Pulsonic Geophysical Ltd of Calgary, Alberta.

Customs agents began their investigation last August when Ivan-Pierre Batinic, a French national and employee of Saxpy at the time, and Kevin Anderson, a software designer from Fremont, California, attempted to cross the Canadian-United States border at Vancouver with approximately \$10,000 in \$100 bills. Anderson and Batinic reportedly received the money from Carlos Julio Williams, a name used on a Guatemalan passport by US citizen Charles McVey. US officials believe McVey was the mastermind behind the technology transfer scheme. He is currently under arrest in Canada, and US officials are seeking his extradition.

A spokesman for the US Customs service says that in addition to arresting Anderson and Batinic, agents also arrested Batinic's brother Stevan, and recovered computer tapes, floppy disks and documents stolen from Saxpy. The men are expected to be charged with violating US export control laws and attempting to transport stolen materials across state lines. There have been press reports that at least one of the group made contact with Soviet officials, including Dr Roald Sagdeev, but customs officials will not confirm this.

Joseph Palca

Australian scientific know-how commercially exploited

Sydney

LIKE the inhabitants of other English-speaking nations, Australians believe their past record shows them to be strong on originality but weak on turning ideas into commercial products.

Now, two inventions — a fast Fourier transform (FFT) VLSI chip and the Sarich engine — have come along to test whether bad old habits can be broken.

The list of inventions on which Australia has failed to capitalize is truly impressive. In the last century, Australia was first with the ice-making plant — but its ice, full of air bubbles, was rejected by consumers in favour of more expensive natural lake ice imported from North America. Next came the light bulb, independently developed a month after Edison's success. More recent failures are the 'black-box' aircraft flight recorder, produced overseas because Australian civil aviation authorities refused to back it, and the Interscan aircraft instrument approach and landing system. Interscan was developed by the Commonwealth Scientific and Industrial Research Organization (CSIRO) Radiophysics Division and chosen as a world standard but no Australian manufacturer could be found. CSIRO collects only a small royalty while the profits from millions of dollars worth of sales go to other countries.

CSIRO Radiophysics Division also developed a programmable computer at the same time as the United States and the United Kingdom. Their 2,000 vacuum tube machine, CSIRAC, was fully operational in 1951, the year the first commercial computer was produced by Ferranti. But according to Dr E.G. Bowen, then head of the division, further development was killed by three reports, two from British scholars recommending that Australia concentrate on agriculture, and a third from Australia concluding that the future belonged to analogue, not digital, computers.

The Radiophysics Division has now come up with the FFT VLSI chip. Its designers claim that it can perform a fast Fourier transform a hundred times faster than the \$150,000 VAX computer and will make tasks like the construction of images by medical scanners possible in real time, rather than after a 15 minute processing delay. According to Dr Bob Frater, the present division chief, the FFT chip has the potential to be to digital signal-processing what the microprocessor is to computing. Determined not to let a third invention slip away, the division is vigorously persuading Australian companies of its commercial potential.

CSIRO's new masters at the Department of Industry, Trade and Commerce are also anxious not to let opportunities disappear abroad. Along with the Western Australian government, they have invested A\$500,000 in a manufacturing feasibility study of an engine developed from the revolutionary design of a West

Australian, Ralph Sarich. It is a lightweight three-cylinder, two-stroke engine using the orbital fuel injection system developed by Sarich for an engine that relied on combustion around the inside of a drum to rotate an internal core. The Sarich two-stroke is claimed to be cheaper to produce, have fewer parts and use less fuel than normal engines. If Australian industry wakes up there is, says a government spokesman, a market for 250,000 engines a year, with potential exports worth A\$1,100 million. **Charles Morgan**

New science prizes awarded to two Italians, one West German

Munich

THE first Italgas prizes have just been awarded in Italy. The recipients — a West German physicist, an Italian energy researcher, and an Italian computer scientist — each received 100 million Italian lire (about \$76,000) from the Italian natural gas company Societa Italiana per il Gas (Italgas) at a ceremony in Turin on 9 October. The money can be used in any way the winner chooses.

The Italgas Prizes, awarded only to scientists in EEC (European Economic Community) countries, are quite lucrative. By comparison, Nobel prizes bring about \$340,000 for each category; the renowned Albert Lasker awards, given in the United States, are worth only \$15,000 divided among the winners.

The prizes will be awarded for at least the next ten years to researchers in three broad areas — physics and chemistry, environmental and energy science, and materials science, communications and information science. Researchers are chosen for outstanding achievement and deep involvement in their fields; significant progress at a young age is given priority, as is research that has to do with natural gas.

This year's winners were Theodor Hänsch, 47, a physicist at the Ludwig-Maximilian University in Munich; Mario Silvestri, 68, a professor of energetics at the Polytechnic of Milan; and Raffaele Meo, 52, a professor of information systems at the Polytechnic of Turin.

Hänsch has distinguished himself with his work in laser spectroscopy. He has developed techniques for very-high-

resolution laser spectroscopy and applied them to atomic hydrogen and other simple atomic systems. By comparing the results of these very precise experiments with theory, Hänsch can test basic laws of physics in ways not possible using other methods. He hopes eventually to measure better values for fundamental constants such as the Rydberg constant or the mass of the electron. Hänsch returned to West Germany in 1986 after spending 15 years at Stanford University.

Silvestri was an Italian pioneer in exploring the peaceful uses of nuclear energy before shifting to research on non-linear thermodynamics. He collaborated from 1965 to 1976 with US, British and Canadian researchers developing heavy-water and boiling-water reactors.

Meo used his expertise in digital signal processing to produce one of the earliest automatic voice recognition systems. He is currently interested in computer architectures and networks, and is also active in research projects sponsored by Consiglio Nazionale delle Ricerche (CNR), Italy's national research council.

Nominations for the prizes are sought from universities, academies of science and private and governmental research institutions in EEC countries. This year's winners were selected from eighty good candidates, according to Italgas.

Italgas initiated the prizes to celebrate its 150th anniversary and is currently at work at a number of other endeavours on behalf of science. Next on the list is the reopening after many years of the historic Academy of Sciences building in Turin on 31 October.

Steven Dickman



Left to right: Professors Raffaele Meo, Mario Silvestri and Theodor Hänsch.

Nationalist environmental problems in Soviet Union

London

THE Central Committee of the Communist Party of the Soviet Union and the Council of Ministers of the USSR have adopted a resolution launching a long-term USSR state programme for environmental protection and the rational use of natural resources. The programme is to be fully worked out by the end of next year. It will deal in particular with the development of waste-free technologies, the re-equipping of existing industrial enterprises and the best siting of new ones, improvement of environmental monitoring service, a drive for better ecological education of management cadres and a public information programme, studies of the cost-efficiency of environmental protection measures, and the monitoring of "particularly crucial" installations such as high dams and mud-slide traps.

The Soviet public, thanks to *glasnost*, is well aware of the need for such measures. Yet, ironically, virtually all the environmental initiatives of the past two years (which the new programme is intended to answer at the All-Union level) have been triggered by nationalist sentiments. Ukrainians have launched protests against the planned Danube-Dnieper canal and the proposed nuclear power station at Chyhyryn. Latvians, Byelorussians and Lithuanians have successfully halted construction work on the planned Daugavpils hydroelectric station. The Tadjik Academy of Sciences says that a new dam (superfluous to the power needs of the Tadjik SSR) in that highly seismic area would prove a 'sword of Damocles' above the republic's capital Dushanbe.

In the wake of the nationalist disturbances in Kazakhstan last autumn, citizens of Alma-Ata have reminded each other that in 1973 the city narrowly escaped being wiped out by a mud-slide because Moscow planners overruled local experts in the choice of a design for mud-retention installations.

The environmental needs of the fifteen union republics sometimes come into conflict. The seven Russian writers collectively known as the Villagers, whose works proclaim traditional Russian values and life-styles, campaigned vigorously for the cancellation of the southward-diversion of Siberia's north-flowing rivers. But to the inhabitants of Soviet Central Asia, the decision to halt work on the diversion was most unwelcome, as without the Siberian water, the Aral Sea, whose feeder-rivers have been dangerously depleted by irrigation works, is in danger of drying up completely. The Uzbek Union of Writers accordingly established a 'public committee' to save the Aral, and

continues to urge that it is impossible to rescue the sea without the introduction of water from elsewhere. (The official Moscow line is that the locals must learn to manage their existing resources more efficiently.)

Soviet environmental conflicts, however, are not simply a matter of local heroes versus Moscow villains. Some local worthies have been reluctant to join in the ecological debate. The campaign of the Georgian writers against the Caucasian Mountain Railway was launched in April by the quasi-dissident Akaki Bakradze, but only after the party secretary of the republic had publicly criticized Georgian failure to follow the example of the Villagers.

In Armenia, environmental lobbying began as early as 1985, with attacks by leading members of the Armenian Academy of Sciences on environmental threats ranging from the emissions of chemical plants to a nuclear spent-fuel dump to serve the entire Caucasus. On the basis of the scientists' findings, in March 1986, 350 Armenian intellectuals signed an open

letter to Mikhail Gorbachev, protesting that environmental pollution in the republic had now reached horrific proportions.

Nothing, however, was done, until in June 1987, the Moscow *Literaturnaya Gazeta* published an article by Armenian journalist Zori Balayan, "Is Armenia on the brink of an ecological disaster?" As a result of this article, an expert commission was sent from Moscow, which, after a two-week stay, produced a package of recommendations for the closure and relocation of chemical plants and the curtailment of the nuclear energy programme virtually identical with that proposed in the 1986 open letter. The local authorities, however, seem unwilling to take action, so that on at least two occasions in the past few weeks protesters, including party officials, have gathered outside the Nairit chemical installations, one of the chief offenders.

Perhaps an even more disturbing case was reported in Kirgizia last month — a cement works 2 km from Lake Issyk Kul, which was closed down after prolonged local lobbying, was opened again after a few hours on the direct orders of the Kirgiz Minister of the Construction Materials Industry and of the provincial party secretary.

Vera Rich

India's electronics companies drive back foreign competition

New Delhi

INDIA is successfully driving back foreign electronics companies from some of its home markets and may soon emerge as an international competitor in its own right.

The first signs that India could produce its own mass-market electronic technology came when the government-owned Semiconductor Complex Ltd of Chandigarh launched a digital wrist-watch in July. More than a million have been sold and the trade in smuggled watches is coming to an end.

Now, successes are coming in two other areas: digital telephone switching equipment and large-scale integrated circuit chips.

Telephone exchanges are being designed at the Centre for Development of Telematics (C-DOT), the creation of Sam Pitroda, an expatriate Indian who was a successful entrepreneur in the United States. Three years after setting up, and with an expenditure of just \$25 million, C-DOT succeeded in developing a family of digital switching systems for which most components can be manufactured locally.

According to Pitroda, the technology is tailored to the needs of developing countries, which lack air conditioning and have power supply problems, dusty environ-

ments and unskilled manpower. The technology is also tailored to India's low telephone density — just 3.5 million telephones for a population of 700 million.

Large-scale integrated circuit chips are already being exported to Far Eastern markets. They come from Semiconductor Complex Ltd, which soon expects to export a million chips a month. Its first products were 5-micron scale — still primitive by the 2-micron standards needed for standard 256K RAMs — but SCL's 200 engineers have subsequently developed 3- and 2-micron technologies in-house. Next year the company will introduce 1.25-micron technology and it has plans for building an electron-beam facility with money from the World Bank so as to tackle sub-micron designs.

SCL's marketing manager, P.S. Mankhad, says the aim is to counter the competition from international giants "with simple and standardized products at low price". Given India's plentiful supply of inexpensive, skilled manpower, "we can compete with the United States and Japan," he says. Demand is expected to rise sharply, and SCL's chairman Mr Virendra Mohan believes India will be exporting chips to the value of a thousand million dollars worth of chips a year in the early 1990s.

K. S. Jayaraman

New lease of life for Lawrence Berkeley Laboratory

Berkeley, California

LAWRENCE Berkeley Laboratory (LBL) is booming into the 1990s. As one of the chosen centres for the Department of Energy (DoE) Human Genome Project, a DoE designated superconductivity research centre and home for the new \$100 million Advanced Light Source synchrotron, the national laboratory is embarking on a wave of activity and growth.

As if to belie its origins as a nuclear physics centre, laboratory director David Shirley has embarked on a reorganization plan giving the life sciences a much more prominent position under Paul Silverman, now appointed associate director. The human genome project should carry LBL's life sciences programme out of its traditional field of radiation studies and into the mainstream of modern biology.

The DoE's decision to pursue a project to map and sequence the human genome has led to questions as to why it has chosen to become involved in a large-scale biology project. But Shirley says that conducting the genome project in a DoE laboratory with DoE money will protect other biomedical research, such as that sponsored by the National Institutes of Health (NIH).

DoE's decision to name LBL as one of two new centres for human genetic studies is nevertheless something of a surprise. Lawrence Livermore National Laboratory had been more frequently mentioned as a participant, largely on the strength of its work on chromosome sorting, an important first step in dividing the genome into manageable subunits.

But John Hearst, of LBL's chemical biodynamics division and a chemistry professor at the University of California at Berkeley, says LBL has unique advantages that make it a logical place for the project. The laboratory has extensive computing facilities and the technological skills that will be needed in producing a physical map and ultimately a sequence of the 3,500 million base pairs that make up the human genome. It is also within walking distance of the university campus where basic biological sciences are winning renewed attention, adding scientific credibility to DoE's efforts in biology.

Silverman says LBL intends to use the new yeast artificial chromosome (YAC) technology to clone large DNA fragments — from 400 to 1,000 kilobases. Restriction enzymes such as NotI, which cut DNA strands at infrequent combinations of bases, have only recently made it possible to obtain such large pieces of DNA. The new technique of pulsed field gel electrophoresis provides a tool for analysing the large fragments. Charles Cantor, one of

the developers of the new technique, is being recruited from Columbia to head LBL's genome project.

Jack Bartley, deputy director of the genome project, acknowledges that there may be pitfalls in pursuing such a combination of new technology, but he says that early investigations seem to confirm that the large cloned fragments can be reliably reproduced. While LBL pursues the YAC technology, Los Alamos and Lawrence Livermore National Laboratories will continue to work on the smaller 40-kilobase cosmid libraries.

LBL has also encouraged other campuses of the University of California system to submit grant applications to DoE for genome work, all of which would be coordinated at Berkeley.

Silverman is reluctant to discuss budget figures until Congress approves DoE's 1988 budget. DoE has requested \$12 million for genome activities in the fiscal year 1988, and there is sufficient confidence that enough of that will make its way to LBL to enable Silverman to begin recruiting staff for the project.

LBL's energy science programme is also growing, as plans get under way for the Advanced Light Source (ALS), which will generate X-ray beams for research and diagnostic uses. It will be the laboratory's first new high-energy facility to be built in over 30 years.

The X-rays generated by the synchrotron will range in energy from 1 to 1,000 eV. Jay Marx, programme director for the project, said the ALS will complement the higher energy 7-GeV synchrotron planned for Argonne National Laboratory. While the Argonne synchrotron will be more suited to solid-state physics, Marx said Berkeley's ALS will be ideal for the imaging of atoms in biological materials, since water is transparent to X-rays with energy in the 40-eV range. Among other planned applications for the ALS are gas-phase chemistry — as, for example, using X-rays to measure extremely small concentrations of gases — and materials science applications such as determining the orientation of catalytic molecules on surfaces.

As a national laboratory facility, Marx said, the ALS will be open to all researchers whose projects pass peer review. The plans allow for the construction of seven beam lines on present budgets, but the laboratory is looking for companies or research groups wishing to build extra beam lines, at a cost of \$0.5-2 million per line. Marx expects the ALS to be a "gold mine for science and technology", adding that every synchrotron source of the handful already in operation

is oversubscribed with users.

The ALS project received \$9.5 million in start-up money from the DoE in fiscal year 1987, and has asked for \$18 million for 1988 and \$30 million for 1989. Construction is due to begin in 1988 and to be completed in 1992.

For the first time in 20 years, other new construction is under way at LBL. Two buildings nearing completion will house the Center for Advanced Materials, whose members have been scattered between the laboratory and the adjacent University of California campus since the centre's establishment in 1983. Plans are in place for the 2-acre building which will house the Advanced Light Source, and more construction is on the horizon, to house the human genome and superconductivity research centers.

The DoE designated Superconductivity Research Center at LBL, which will study thin film superconductors, is still in the planning stages, without a building or a guaranteed budget. Director Shirley says that the LBL's major worry on all three projects is a concern shared by all US science: whether Congress will commit the funds these projects require.

Marcia Barinaga & Joseph Palca

India at last gets US supercomputer

New Delhi

AFTER two years of negotiations, the United States has agreed to export a supercomputer to India for use in monsoon research. An agreement was signed last month, a few days before Prime Minister Rajiv Gandhi's scheduled meeting with President Reagan in Washington.

Although the president had approved the sale of a supercomputer to India in 1985, the negotiations were held up because of differences over safeguard conditions and the type of supercomputer to be supplied. The agreement signed in New Delhi by John Gunther Dean, the US ambassador, and K.P.S. Menon, the Indian Foreign Secretary, does not specify the model. India had requested a Cray XMP-24, the type sold to US allies. According to a Foreign Office spokesman, the United States has agreed to supply any one of the less advanced versions — the Cray XMP-1, XMP-14 or CDC-205 — but not the twin processor XMP-24.

The Indian Department of Science and Technology is likely to choose the XMP-14. The Indian government is said to have sought assurances from the United States that the computer will be upgraded later. India has also agreed to safeguards and curbs on the use made of the computer, but no details have been published.

K.S. Jayaraman

Boycotting South Africa

SIR—In response to J.G. Wilson (*Nature* **328**, 288; 1987), we should like to present our case against a total academic boycott of our scientists.

We have no doubts that trade and sports boycotts against South Africa have been effective; any government moves towards reform can be attributed to such pressures. That is why academics elsewhere are seeking to produce similar effects by boycotting our universities, as in disallowing South African participation in international conferences, a reluctance to attend conferences in this country and the rejection of South African contributions by international journals.

It is interesting that, largely as a result of these measures, many South African academics have become so conscious — and so defensive — of academic freedom. These same people have, however, remained conspicuously silent about the erosion of freedom when the infringements have involved members of the political left. Such infringements occur daily. The more spectacular examples include the prolonged detention of our staff and students, occupation of the campus by heavily armed police, banning of meetings, tear-gassing and assault of staff and students by police and the destruction of property. But far more common, and perhaps more insidious, are the instances of staff being ostracized by colleagues, the withdrawal of work permits by the authorities, the threat of deportation, the opening of university mail and the tapping of telephones.

All is not, therefore, as Wilson sees it. Far from "living comfortably in South Africa", many of our liberal academics are suffering from oppression. To oppose apartheid in Britain is easy and is even encouraged; to oppose it here is dangerous. How would Wilson survive 13 months' detention in prison without trial, as has our colleague Raymond Suttner? The price we pay for opposing the government in this country can be very high indeed. It can be only a matter of time before a student is killed on campus.

Is a total academic boycott likely to advance the struggle against apartheid in any way? We remain unconvinced. Academic isolation will not be immediately destructive, but it will have dire consequences in the long term. Traditions of academic excellence are not built overnight, and, once destroyed, are almost impossible to rebuild. It is clear that the present government will not relinquish power until it is forced by bankruptcy to do so. Our economic recovery thereafter will depend on the continued existence of a sophisticated system of education and training. In stifling the monster of apartheid, we must take care to protect the

embryonic nation that is struggling to be born.

If our educational institutions are destroyed now because of the isolation of their academics, no subsequent government will be able to raise the country from its inevitable plunge into under-development and exploitation. By all means, let us continue to fight for the equal allocation of South Africa's resources to all its people, but let us also ensure that there will be something worth sharing, once this government has ceased to exist.

A total boycott of the South African education system is as wrong as no boycott at all. What we suggest is a system to identify opponents of apartheid within South Africa based on a signed declaration that echoes the aspirations of the Freedom Charter:

All the cultural treasures of mankind shall be open to all, by free exchange of books, ideas and contact with other lands;

Education shall be free, compulsory, universal and equal for all children.

Higher education and technical training shall be opened to all by means of state allowances and scholarships awarded on the basis of merit.

Let all South African academics who wish to be considered a part of the international community and those overseas academics who wish to visit here and speak on our campus be invited to sign such a declaration. This implies the implementation of selective sanctions against those who support the principles of apartheid. It is only by such action that academic opposition to apartheid both inside and outside this country can be united.

In this way we should, through academic responsibility, ensure that the liberal universities retain their standards of academic excellence, that they remain members of the international community and that they resist intimidation from all who would break down their liberal ideals.

R.J. RAYNER
J.C. MASTERS

*University of the Witwatersrand,
Witwatersrand 2050,
South Africa*

SIR—The clear vision, unquestioning faith, absolute judgement and fervent determination that J.G. Wilson (*Nature* **328**, 288; 1987) brings to bear in his discussion of the academic boycott of South Africa are exactly those attributes that characterize the Nationalist Party and its relentless misgovernment of South Africa.

Neither the nationalists (of whatever 'colour') nor Wilson seem to realize that the present political crisis (as it relates to racism) is destined to be overrun by the real problems of population growth. While the various factions invest their effort in

the battle for political supremacy, they appear to be oblivious of the forces that could reduce their inheritance to a wasteland.

The population of South Africa is expected to increase from the present 30 million to 80 million by 2020. So, for the foreseeable future, the population profile will be dominated by children who do not contribute to the country's work-force but who place an ever-increasing load on agriculture, water supply, health, education and technical services. Already more than half of the population of South Africa is below the age of 20. South Africa must maintain a real growth rate of 4.5 per cent merely to provide employment opportunities to meet the population growth. This is hardly the time to persuade trained people and entrepreneurs to leave the country, or to ostracize those who remain. Even less is it the time to destroy (as opposed to transform) the economy required to support new lives as well as to provide for a new political dispensation with a generally enhanced standard of living.

Under any other circumstances, there would be little doubt about the solution to a problem of this kind. We would recommend that skills, education and agriculture be nurtured. We would recommend the development of a robust local industry with a significant export market. We would encourage entrepreneurs and investment. Instead, both the Nationalist Party and the international community are blinded by the issues of racism. The Nationalist Party is prepared, if necessary, to spend its last years dragging southern Africa into the abyss of war. The international community, intent on vengeance, is determined to see the downfall of 3 million white Nationalists at the cost of decimating the whole infrastructure of that same subcontinent. Their position is usually justified on the grounds that, hopefully, the devastation need only be transient. Unfortunately, phoenix-like revival of fortunes has not been a characteristic of the post-colonial developing countries. Wilson and his like would do well to be more circumspect before encouraging 30 million people onto a course of self-destruction.

I suggest that the only strategy that could save the southern African region from an irreversible slide into economic and social oblivion is determined and selective foreign investment in black business and education. Racism can be conquered by education and economics, but not by wars.

RODNEY J. DOUGLAS

*104c Banbury Road,
Oxford OX2 6JX, UK*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

New ways with reptating polymers

A surprisingly simple model of polymer molecules moving in a network representing the constraints of others like themselves may be a pointer to a simpler theory of macromolecules in solution.

THE simplest thing to say about the thermodynamics of polymer solutions is that nothing is particularly simple. This, no doubt, is one of the reasons why attempts to calculate the properties of polymer systems seem ever to be striving for still simpler models. One of the neatest so far must be that of Michael Rubinstein, from the Eastman Kodak Company at Rochester, New York, who has produced a model for the movement of a polymer molecule that resembles in many ways the movement of the pieces on a backgammon board (*Phys. Rev. Lett.* **59**, 1946; 1987).

The essence of the problem is that polymer chains can become entangled with each other. In extreme cases, they can even become knotted together, simulating molecules twice as big, or even several times as big, as the starting material. But even when this has not happened, the movement of individual molecules will be constrained by the presence of neighbouring molecules.

What this means in practice is that not all the space a polymer molecule occupies is freely accessible to it. Some parts of the molecule may be relatively free to assume all possible configurations, but others will be constrained as if at bottlenecks — there may be points, or regions of space represented perhaps by encircling loops of other molecules, through which they must pass.

The calculation of the movement of a polymer molecule, or even the mere enumeration of the configurations accessible to it, is a little like estimating the ease with which a long snake can thread its way through a series of bottlenecks along its path. But snakes usually know where they are heading, whereas the motion of polymer molecules in solution is strictly statistical. There is nothing to prevent the movement of polymer molecules being impeded by their own perverse tendency to bunch themselves together in ways that make their passage through bottlenecks that much more difficult.

Plainly all calculations of such a state of affairs must be simplifications. The ground rules are essentially those due to S.F. Edwards (now Sir Sam Edwards of the Cavendish Laboratory, Cambridge) in the 1960s: think of a single polymer molecule in a three-dimensional liquid as being confined within a tube of variable diameter, representing the external constraints of other molecules. Then set out

to calculate the chance that the molecule will move as a whole along the tube in one direction or the other, allowing for the way in which, by chance, different lengths of polymer may be bunched together in successive sections of the tube. This phenomenon, called 'reptation', is what happens when real snakes move.

One of the standard simplifications of the problem of movement by reptation is to suppose that the space accessible to a molecule is defined by some kind of regular lattice. In two dimensions, the ground rules would have it that the backbone of a molecule should not be allowed to cross a lattice point. In three dimensions, the corresponding rule is that polymer tracks entering a cell through one face cannot cross any of the edges bounding that face.

The essence of Rubinstein's further simplification of the problem is his concept of 'reptons', or sections along the length of the molecule (which ideally should be of roughly equal length). Then it is possible to represent a configuration of the molecule as a statement of the number of reptons to be found in successive cells along the prescribed track of the molecule. Naturally, in such a description, there can be no gaps along the track, for that would mean that the backbone of the molecule is broken.

The movement of molecules is now stripped down to, and perhaps beyond, its bare essentials. Rubinstein's recipe is to calculate the statistical properties of these collections of bunched reptons in successive cells under the assumption that each of them is free to move by random walking in either direction subject only to two constraints, which are as follows: (1) the configuration can never correspond to a broken chain, with an empty cell separating two clusters, and (2) the chance that a repton will move into an empty contiguous cell is greater than that it will move into a contiguous occupied cell by the factor $(z - 1)$, where z is the number of faces of each cell, or its coordination number.

The statistical mechanics of this model are breathtakingly simple. With N reptons per molecule, the calculation of the number of configurations in which the reptons are distributed over only K contiguous cells (where $K \leq N$) is simply the number of ways in which $N - K$ objects can be distributed among K sites. From this combinatorial way of putting things, it

quickly emerges that the average value of K , the cluster size, is proportional to the total number of reptons N , which is what everybody expects.

The surprise is that this simple model also yields an estimate of the diffusion constant of the molecule. The first step is to calculate the chance that a jump of the chain from one repton distribution to another will succeed, which is a simple if tricky piece of probability calculation. Taking the diffusion constant as measured by the rate at which the centre of mass of the system moves along its tube, the simple result (in three dimensions) is that the diffusion constant is inversely proportional to the square of the molecular mass (which is the correct result).

This by itself is well-known. Rubinstein's special claim is that his simple model also yields a basis for calculating the relaxation time of a polymer molecule, which is in turn related to the experimental quantity of polymer solutions, the viscosity. The physics underlying this calculation is also childishly simple — to tackle the question of the rate at which a bunched-up polymer molecule will move when one of its ends is tugged (biasing the directions of likely random walks near the pulled end of the chain).

Rubinstein is exercised by this point after noting that there is a glaring discrepancy between previous predictions of the relationship between longest relaxation time of a polymer molecule and its molecular weight (a cube-law dependence) and empirical values (which yield a power-law dependence with an exponent more like 3.4).

As always in these circumstances, the argument concludes with numerical simulations. Rubinstein calculates from his simple model values of the relaxation time or viscosity of longish polymer molecules and concludes that there is no reason why the exponent of the power-law dependence on molecular weight should not be anywhere between 3.22 and 3.52, depending on the value of z . By itself, that proves nothing. The model may be too simple to be a good guide to reality. And who, in any case, can tell what may happen in the real world of polymer molecules, where z may change from one cell along the backbone of the molecule to the next. Yet this is a long way to have come with such a stripped-down model of polymer solutions.

John Maddox

Cognition

Connections and controversy

P.N. Johnson-Laird

THE brain appears to be a parallel computer constructed out of slow and expendable processing units. Its power presumably derives from the way in which these units — the neurons — are wired up. But how can relatively simple processors be connected to carry out interesting computations? Recent theoretical work in cognitive science solves a long-standing problem in the theory of neural networks and throws light on how such computational systems can learn.

The fundamental thesis of the theory of computation is that any computable function can be calculated by a very simple serial automaton (a so-called Turing machine) equipped with unlimited memory. Hence, cognitive scientists have for many years modelled mental operations in programs written in high-level languages, such as LISP, that run on conventional digital computers. Yet they have always known that neurons do not operate according to LISP-like procedures, nor to the instruction set, say, of the 68000 chip. They have made various attempts to specify the brain's computational architecture in the form of networks that connect up many elementary processing units resembling idealized neurons. The difficulty of this exercise can be illustrated by the fate of one such device, the so-called 'perceptron'.

The perceptron consists of a simple retina connected to a layer of input units, which in turn are connected to output units. It can learn to recognize patterns by adjusting the strengths of connections between its two layers of units using a simple regimen based only on information about the correctness of its performance. But certain patterns (for example, exclusive disjunctions of the form A or else B, but not both) are beyond its competence: it cannot identify them, let alone learn to identify them². Its problem, like the problem with traditional associative theories, is that it relies on direct links from input to output units: it makes no use of internal representations.

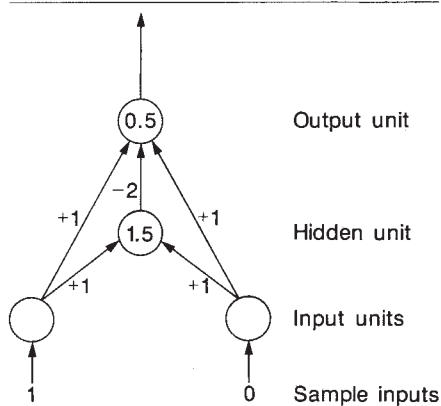
Recognition

A network for recognizing exclusive disjunctions and other complex patterns needs such representations, which can be introduced by way of one or more layers of units hidden between the input and output layers (see figure). Unfortunately, no-one knew how to devise a method of learning for networks with such 'hidden' layers. The strength of each connection between units has to be changed to improve the performance of the network as a whole,

and the critical problem was to determine how much responsibility to assign to a hidden unit in adjusting its strength of connection to an output unit.

This problem has now at last been solved, and there are several methods of learning that have been successfully implemented in networks with hidden units^{3,4}. In the currently most successful method⁴, the 'back-propagation' of error, a network is first used to propagate activation up from the input, through the hidden units, to the output units. The output activation is then compared with the required target activation, and finally the strengths of connection between units are adjusted so as to reduce any discrepancy. This cycle is repeated perhaps for thousands of times until the strengths of connection have stabilized. In general, the contribution that a hidden unit makes to the overall error depends on its level of activation and its strengths of connection to each unit at the next level up (and those units' contribution to error). Hence, the learning algorithm can work back down the hierarchy of units adjusting the strengths of connection appropriately. The algorithm is simple, though excessively time-consuming when it is executed on conventional serial computers.

Such algorithms will often develop



A network with a 'hidden' unit that recognizes when one or other but not both of its inputs is active, that is, an exclusive disjunction. Each unit transmits its activation $(+1) \times$ the strength of its connection to a unit (shown beside the arrow). Each unit fires when its threshold (shown within the unit) is exceeded by the sum of its inputs. Given the sample input shown in the figure, the output unit receives an activation of $+1$ from the left-hand input unit, which exceeds its threshold of 0.5 , and so it fires. If the sample input were 1 and 1 , both input units would fire and thereby trigger the hidden unit, which in turn would inhibit the output unit from firing. Without hidden units, no such network could recognize an exclusive disjunction.

associations between input and output patterns in a way that is distributed over many processing units: each unit participates a little in all the associations, and each association is distributed over all the units. Thus, a series of associations will be merged together into a single composite like a set of superimposed photographs. The network, in effect, captures their similarities without having to represent them in an explicitly structured principle. It depends on parallel distributed processing — a special case of the 'connectionist' architecture in which computation is carried out solely on the basis of the activation of units and the strengths of connection between them. There are many such architectures in theory, and currently much work is concerned with exploring their implications.

Advantages

The chief theoretical advantage of connectionist schemes is that they deliver as emergent properties several phenomena that are known to occur psychologically. For example, the recall of a memory becomes a matter not of search for the contents of a particular pigeon-hole, but rather of an attempt to reconstruct a complete representation of an event from a fragmentary input. Human memory works in the same way — there is no clear-cut boundary between recall and reconstruction. Another theoretical advantage is that the computational architecture is closer to the way in which the brain is likely to work, though it is certainly not wired up in a way that is required by current connectionist proposals. Despite these advantages, however, connectionism has its critics and is the centre of intense controversy.

At the heart of the matter is the issue of the emergent properties of connectionist networks, which has come to a head over a particular study⁵ of children acquiring their native tongue. Children learning English reach a stage in which they correctly use irregular past-tense verbs, such as 'came' and 'went'. Subsequently, they appear to follow the principle that all past tenses are formed by adding the appropriate '-ed' affix (which calls for three distinct phonological forms). At first, they even apply this principle to irregular verbs, and they will say 'camed'. Thus, linguists and psycholinguists, in common with other cognitive scientists, have argued that the mind contains unconscious principles with an explicit structure, such as the principle governing the past tense. These principles are obviously not available to a child's conscious introspection, but if they could be examined, they would be similar to instructions in a high-level programming language. But is it really necessary to suppose that such explicit principles exist in the mind? Perhaps not. Perhaps the behaviour that

appears to conform to them is merely an emergent property of parallel distributed processing.

Such a scheme for modelling children's acquisition of the past tense has been proposed⁶ by Rumelhart and McClelland (although it does not use hidden units), and it gives a superficially plausible account of the phenomena. But a detailed critical analysis⁷ carried out by Pinker and Prince shows that the network's performance diverges significantly from the psychological facts, and that it is in principle unable to account for some of them. One important discrepancy concerns the statistics of the utterances that children hear, as once a network has stabilized, the only force for change will be a change in the patterns of information impinging on it. Network theorists accordingly assume that adults start to use inflected past tenses only after a stage in which the irregular verbs predominate. The evidence suggests that there is no such change in usage. Thus, Pinker and Prince treat the child's linguistic system as an active formulator of explicit principles, which at the earlier stage has erroneously inferred that irregular past forms, such as 'came', are complete words in themselves. When the system postulates the inflectional principle, it is applied to the irregular forms as well.

Unlike many psychological controversies, this one is likely to be fruitful if only because it has exposed some clear theoretical options. One such option is that high-level principles are paramount, and that parallel processes are merely the brain's low-level language into which all mental life must ultimately be translated: high-level principles are, as it were, 'compiled' into low-level parallel processes. At the other extreme, another option is that the mind contains no high-level principles: there is only parallel distributed processing from which all behaviour emerges. Hence, theoretical descriptions of such behaviour in terms of high-level principles or LISP-like instructions are at best misleading approximations. The drawback of this latter view, however, is that human beings can certainly entertain high-level principles in their conscious thinking — science itself would otherwise be impossible. Even if

one were to argue that consciousness is an epiphenomenon, the form of its contents still stands in need of explanation.

The moral appears to be that the mind must, at a certain level of organization, represent explicit individuals, properties, relations and principles. This conclusion is entirely compatible with the existence of mental phenomena that are emergent properties of lower levels of computational architecture. The profound question is to determine the boundary between

such levels, because high-level explicit principles, if Pinker and Prince are correct, are not merely to be found in the conscious mind, but articulate our unconscious knowledge of language. The answer to this question is likely to require a better understanding of the properties of connectionist systems, and of how explicit principles can be compiled into them. □

P.N. Johnson-Laird is Assistant Director of the MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK.

Solid-state chemistry

Aluminium avoidance avoided

J. Stebbins

OXYGEN, silicon and aluminium are, respectively, the three most abundant elements in the Earth's crust. SiO_4 and AlO_4 tetrahedra, connected at their corners by shared oxygens, make up the rigid framework of the most common crustal minerals, as well as most of the phases in oxide ceramics and catalysts. The exact location of the silicon and aluminium atoms in a crystal structure influences its symmetry, its physical and chemical properties and, in particular, its 'third-law' or configurational entropy. ^{29}Si nuclear magnetic resonance (NMR) spectroscopy is now the tool of choice for sorting out such Si–Al distributions. On page 56 of this issue, Klinowski *et al.*¹ report the discovery, by this technique, of a structure that confounds one of the most commonly cited assumptions about the ordering of silicon and aluminium in minerals.

Information about this ordering is essential in constraining thermodynamics and in predicting phase equilibria, and of course a complete picture of a crystal structure is required for a quantitative understanding of its physics. In a mineral such as albite ($\text{NaAlSi}_3\text{O}_8$), for example, the difference in entropy and enthalpy between structures with an ordered or disordered Si–Al distribution can lead to differences of hundreds of degrees in the temperatures of equilibria with other solids. The very reason that aluminium and silicon substitute freely for each other in minerals, their proximity in the periodic table, has often frustrated attempts to distinguish their crystallographic positions.

Because the two elements scatter X-rays in so similar a manner, diffraction methods can only produce a picture of the distances between (undistinguished) tetrahedral cations and their surrounding oxygens. As Al^{3+} is slightly larger than Si^{4+} , site occupancies can be derived, but these are averaged over all of the equivalent sites in the crystal. Neutron diffraction can distinguish aluminium and

silicon more directly, but still reveals an average structure.

An NMR spectrum is primarily controlled by the local chemical environment, rather than by the long-range periodicity, so that NMR is an excellent complement to diffraction techniques. In the inorganic realm, probably the most useful application has been to study silicon in aluminosilicates, particularly in the industrially important zeolites. Because the ^{29}Si resonant frequency (quantified as the chemical shift) is very sensitive to the chemical type of next-nearest neighbours, the precise local distribution of tetrahedral cations can often be derived, in some cases for the first time.

In 1953, Walter Lowenstein² pointed out that few natural minerals have more aluminium than silicon cations in tetrahedral sites and that links between AlO_4 tetrahedra are rare or absent. This finding is general enough to have become called Lowenstein's rule, or the aluminium avoidance principle. Its most obvious prediction is that in minerals with equal proportions of tetrahedral aluminium and silicon, the cation distribution should be fully ordered, with strict alternation of silicon and aluminium atoms. In structures with a greater Si/Al ratio, some disordering can occur, but Al–O–Al bonds should not be found. In materials which do contain a high fraction of tetrahedral aluminium, such as CaAl_2O_6 , the rule cannot be maintained, of course. Theoretical treatments^{3,4} have shown that aluminium avoidance need not always be exact, but the principle has remained a useful guide in estimates of configurational entropy.

Recent ^{29}Si NMR studies have found many vindications of Lowenstein, and some violations. In the nepheline-group minerals (roughly $(\text{Na,K})\text{AlSiO}_4$), for example, there is strict Si–Al short-range ordering even in very high-temperature phases⁵, although X-ray diffraction indicates long-range disorder. In many zeolites (usually sodium–aluminium silicates), the rule is generally obeyed. On the other

1. Rosenblatt, F. *Principles of Neurodynamics* (Spartan, New York, 1982).

2. Minsky, M.L. & Papert, S. *Perceptrons* (MIT, Cambridge, 1969).

3. Hinton, G.E. & Sejnowski, T.J. in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1 Foundations* (eds Rumelhart, D.E. *et al.*) 282–317 (MIT, Cambridge, 1986).

4. Rumelhart, D.E., Hinton, G.E. & Williams, R.J. *Nature* **323**, 533–536 (1986).

5. Brown, R. *A First Language: The Early Stages* (Harvard University Press, Cambridge, 1973).

6. Rumelhart, J.L. & McClelland, J.L. in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2 Psychological and Biological Models* (eds McClelland, J.L. *et al.*) 216–271 (MIT, Cambridge, 1986).

7. Pinker, S. & Prince, A. *On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition* (Center for Cognitive Science, MIT, 1987).

hand, metastable cordierite ($\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_{10}$), synthesized at high temperature from a melt, has abundant Al—O—Al linkages that are removed as the crystal is annealed⁶. X-ray data on a metastable hexagonal polymorph of anorthite ($\text{CaAl}_2\text{Si}_2\text{O}_8$) suggest that it, too, is disordered. A tentative pattern is suggested here: structures with interstitial cations that are smaller and more highly charged (giving a higher field strength) may have a greater tendency to form Al—O—Al links to enable greater localization of negative charge.

The intriguing report of Klinowski *et al.* in this issue¹ may be the first definitive report of true local Si—Al disorder producing abundant violations of aluminium avoidance in a sodium aluminosilicate mineral with a Si/Al ratio of about 1. The authors also point out the contrast between the disordered, synthetic phase

and its ordered natural counterpart. It is likely that the synthetic phase is formed far from equilibrium, and thus is highly metastable. The energetic implications for even more disordered, less stable phases could be significant. Those of us studying glass must be careful when assuming that Lowenstein's rule for equilibrium crystals also applies to their amorphous counterparts. □

1. Klinowski, J., Carr, W.W., Tarling, S.E. & Barnes, P. *Nature* **330**, 56–58 (1987).
2. Lowenstein, W. *Am. Mineral.* **39**, 92–96 (1953).
3. Barrar, R.M. & Klinowski, J. *Phil. Trans. R. Soc. A* **285**, 637–676 (1977).
4. DeJong, B.H.W.S., Schramm, C.M. & Parziale, V.E. *Geochim. cosmochim. Acta* **47**, 1223–1236 (1983).
5. Stebbins, J.F., Murdoch, J.B., Carmichael, I.S.E. & Pines, A. *Phys. Chem. Miner.* **13**, 371–381 (1986).
6. Putnis, A., Fyfe, C.A. & Gobbi, G.C. *Phys. Chem. Miner.* **12**, 211–216 (1985).

J. Stebbins is in the Department of Geology, Stanford University, Stanford, California 94305, USA.

Archaeology

Getting blood from stone tools

Paul G. Bahn

NEW techniques developed by the Canadian scientist Thomas Loy have led to the discovery of Neanderthal blood on a stone tool from a site in Iraq, and to the accurate identification of the animal species on which other stone tools were used. Extracting blood from stones is likely to be of great importance in terms of new archaeological knowledge, and a massive amount of work remains to be done on the millions of stone tools lying in the museums of the world. These tools may hold the kind of information which early archaeologists could only dream of obtaining.

The archaeological world was both startled and excited in 1983 when Loy, of the British Columbia Provincial Museum, Canada, claimed¹ to have found blood residues surviving on the edges of ancient stone knives. After use of a tool, blood dries and fixes quickly, although some

seeps into the surrounding soil when the object lies buried. Providing that the right combination of temperature, moisture and acidity is present, the haemoglobin survives intact, and can therefore be crystallized out and analysed, assuming that the tool is not cleaned too well after excavation. The shape of the crystals of haemoglobin varies between animal species, and provides a means of identifying on which animal a tool was used.

Loy studied 104 tools of chert, basalt and obsidian from open-air sites in coastal British Columbia, with dates from 1,000 to 6,000 years ago, and identified haemoglobin from animals such as moose, caribou, grizzly bear and sea-lion. He also obtained radiocarbon dates from some blood residues using accelerator mass spectrometry, which requires only minute samples for analysis.

An artefact from Strawberry Bluff, a site in northern British Columbia, produced a corrected date of $1,010 \pm 90$ before present (BP), which matches a date from the charcoal of a nearby hearth ($1,060 \pm 160$ BP). The blood turned out to be that of a snowshoe hare (*Lepus americanus*). A bifacial knife from Toad River Canyon, a site in the same region, had residues of both caribou and human blood, and produced a date of $2,180 \pm 160$ BP, which fits the style of this tool (the site is now destroyed and was never dated).

Loy has now extended and improved the technique². He finds that blood residues can survive on tools for at least 100,000 years: three tools from Barda Balka, north-east Iraq, had deposits of mammalian (probably ruminant) blood, and this open-air site has been dated

geologically to between 75,000 and 125,000 years ago. Other examples have been found on tools of lesser age from Belgium (20,000 BP); Jarmo, Iran (8,000–10,000 BP); and Nevada (6,000 BP) as well as those from northern Canada, indicating that such residues can survive in a wide range of environments. It is therefore possible that they will eventually be found on tools of far greater antiquity.

Loy's original method of extracting and crystallizing the haemoglobin in the residues has now been supplanted by the electrophoretic separation of all blood proteins, a method in which protein molecules are separated by their different mobility and migrations in a gel to which an electric field is applied. The pattern of blood separation is compared with control samples from different animals, and the species can thus be identified. Where the identification points to human blood, Loy uses an immunological testing procedure; of the 25 tools analysed in this way to date, 18 have positive reactions for the presence of human immunoglobulin. In one case (a tool from Barda Balka), the blood is almost certainly Neanderthal.

This result by no means necessarily indicates the existence of cannibalism or sacrifice. Every flint-knapper or excavator of sites with stone tools knows only too well how easy it is to slice your finger open on a sharp edge. In more recent cases, however, the blood is directly associated with the cause of death, as in the residue on an arrowhead found in the rib cage of a young female skeleton of about 1,400 years ago. Elsewhere, a burial ritual seems to be involved, as in the bifacial knife found³ with some dismembered bodies from a mass grave in eastern Canada of the fourteenth century AD.

The possibilities for analysis of human blood residues are enormous, ranging from research on the history of specific diseases to work on genetics and human evolution, and isotope analysis for reconstruction of past environments and diet⁴. Advances have recently been made in the study of use-wear on ancient tools, and of residues on their edges such as fragments of hair, feathers or plants, all of which can help to explain their function. The discovery of blood residues not only pinpoints the species on which the tool was last used, but can also produce an apparently accurate date for that use — a great step forward, as stone tools could previously be dated only by style or by stratigraphic position. □

1. Loy, T.H. *Science* **220**, 1269 (1983).
2. Loy, T.H. in *Archaeometry: Further Australasian Studies* (eds Ambrose, W.R. & Mummery, J.M.J.) 57 (Australian National University, Canberra, 1987).
3. Loy, T.H. *Can. J. Archaeol.* (in the press).
4. Sealy, J. *Stable Carbon Isotopes and Prehistoric Diets in the South-Western Cape Province, South Africa* (BAR Int. Series 293, Oxford, 1986).

Paul G. Bahn is a freelance writer of archaeology at 428 Anlaby Road, Hull HU3 6QP, UK.

100 years ago

SYNTHESIS OF GLUCOSE

ANOTHER important acquisition to our store of knowledge has recently been made. Glucose has been artificially prepared by Drs Emil Fischer and Julius Tafel in the chemical laboratory of the University of Würzburg. This happy achievement has long been looked forward to, and cannot fail to give deep satisfaction in chemical circles all over the world. Not only has the sugar itself been actually prepared, but considerable light has been thrown upon that much-discussed question — the constitution of sugars. A remarkable attribute of this artificial sugar is that it is incapable of rotating a beam of polarized light. In preparing a glucose of the composition $\text{C}_6\text{H}_{12}\text{O}_6$ the probability is that both dextro and laevo are simultaneously formed, and thus neutralize each other, producing a totally inactive mixture. From *Nature* **37**, 7; 3 November 1887.

Molecular biophysics

What about protein polarity?

Arieh Warshel

ONE of the key open problems in the field of genetic engineering is the correlation of the amino-acid sequence and the three-dimensional structure of proteins with their functions. Electrostatic energies have been suspected for some time^{1,2} to be a major factor in such correlation. Yet attempts to address the problem quantitatively have been rather limited³. Most studies use entirely qualitative calculations and graphical displays of electrostatic potentials (evaluated with the vacuum dielectric constant). But genetic-engineering techniques have changed the situation in a dramatic way. It is now possible to examine theoretical predictions directly by experimentally creating or removing charged residues, and even changing the hydrogen-bonding pattern around such residues. This challenges theoretical models in two ways: first, to reproduce the effective interactions between charges; and second, to reproduce the energy of interaction between charged groups and their local environment. The first and somewhat simpler challenge is addressed in two papers in this issue^{3,4} by Sternberg *et al.* on page 86 and by Gilson and Honig on page 84.

The translation of the three-dimensional structure to electrostatic free energies is far from simple. In particular, it is not obvious how to evaluate the interaction between charges in the complicated environment of the active site of a protein. It is not clear whether such charges should be treated as if they are in a hydrophobic (oil-like) or in a hydrophilic (water-like) environment, or if neither case can adequately describe the actual environment. The two papers in this issue use similar methods based on treating the water molecules around the protein as a continuum with a high dielectric constant and treating the protein itself as a continuum with a low dielectric constant. The traditional problems associated with the complicated shape of the protein (which cannot be approximated by a simple sphere) are overcome by dividing the solvent into a fine grid of volume elements and solving numerically the relevant macroscopic equation.

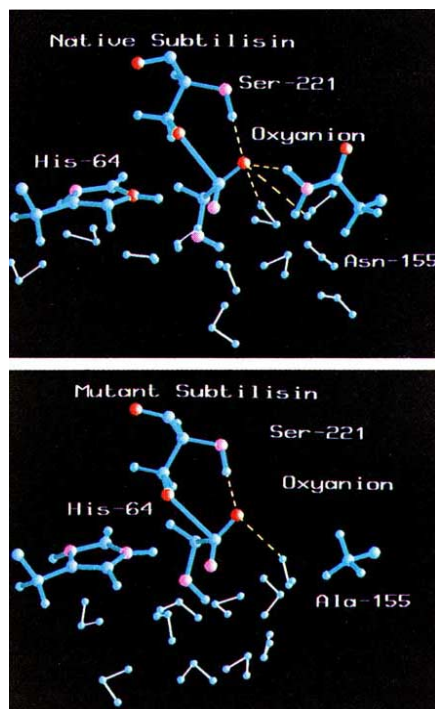
Encouraging results

Although the solvent around the protein is not a macroscopic continuum, the results obtained by such models seem quite encouraging (even though in some cases⁴ the relative errors are rather large). However, the reported calculations are confined to weak (less than 1 kcal mol⁻¹), long-range electrostatic interactions

between groups on the protein surface. Whether the same method could work for strong short-range interactions inside the protein is still not clear (see below). Yet one very instructive conclusion emerges from the work of Sternberg *et al.*³ and from earlier studies²; the interaction between charges in proteins can be estimated using a simple Coulomb's law of the form

$$\Delta G = 332 Q_1 Q_2 / R \epsilon(R) \quad (1)$$

where ΔG is in kcal mol⁻¹, R is the distance between the charges (Q_1 and Q_2), and $\epsilon(R)$ is an effective dielectric constant that represents the effect of the protein and water surrounding the two charges. The surprising result is that the effective dielectric constant ϵ is large, being between 60 and 30 in most cases. This finding, which might shock those who view proteins as hydrophobic spheres, is not entirely due to the surrounding water molecules but to the fact that the protein itself has very polar components that can



Stabilization of the oxyanion transition state in the catalytic reaction of the native (Asn 155) and mutant (Ala 155) subtilisin. The mutation changes the local polarity and reduces the electrostatic stabilization of the transition state by about 3.7 kcal mol⁻¹. Calculations of this change in 'solvation' free energy gave around 4.1 kcal mol⁻¹. Because the observed effect is caused by the local hydrogen bonding, it cannot be reproduced by continuum models. (From the simulations reported in ref. 8.)

'solvate' charges quite effectively² (see below).

How general is the application of equation (1), and can it really be a substitute for microscopic modelling²? The answer usually depends on the specific problem. The use of this equation for interactions between charges that are situated far apart indeed captures the main physics of small interactions between the charges. But weak long-range electrostatic interactions are not so important. The cases that really matter involve short-range interactions (R smaller than 4 Å) where the local polarity stabilizes ion pairs much more effectively than would be predicted using equation (1) and the continuum models. Functionally important ion pairs with electrostatic interactions of more than 5 kcal mol⁻¹ can now be identified with certainty.

Challenge

The reproduction of such interactions is a serious challenge for the available theoretical models. It seems clear that only models that take into account both the polarity of the water and of the protein could reproduce the observed energetics. The effect of the bulk water and to a far lesser extent the bound water molecules is taken into account by the macroscopic models reported in both papers in this issue^{3,4}, but the treatment of the protein polarity requires models of a more microscopic nature².

An instructive test case has been reported recently in a study of aspartate aminotransferase⁵, where a very strong ion pair (Asp⁻--Arg⁺) is inverted by genetic engineering to a pair of reversed polarity (Arg⁺--Asp⁻). The interaction energy of the (−+) system is larger by at least 2.7 kcal mol⁻¹ than that of the (+−) system (the difference at the transition state is more than 6 kcal mol⁻¹). Apparently the relevant dielectric constant changes from about 13 to 80 by simply changing the polarity of the ion pair. Obviously, this result cannot be reproduced by equation (1), nor by macroscopic models that do not take into account the protein polarity. Reproducing this polarity reversal experiment and other cases of very stable ion pairs could provide a way to discriminate between different electrostatic models.

Although studies of charge-charge interactions are instructive, more unique test cases are associated with the second part of the electrostatic energy, the so-called 'self' energy or the energy of forming the charge in the given environment. This factor is frequently overlooked because for many people electrostatic energy is simply the attraction between positive and negative charges. Yet an isolated ion possesses an energy of its own; it likes to be in polar environments and dislikes nonpolar ones. This fact can be expressed on the macroscopic level by

Born's formula:

$$\Delta G = -166(Q^2/a)(1-1/\epsilon_B) \quad (2)$$

where ΔG is given in kcal mol⁻¹, ϵ_B is the relevant dielectric constant for this solvation process and a is the effective radius of the given charges. It follows from equation (2) that polar sites with large ϵ_B give much larger solvation energy than non-polar sites. Proteins can exploit this point to 'solvate' isolated ions as much as or more than water by using their dipolar components (such as N-H and C=O dipoles). This point seems to surprise those who expect to find ion pairs in the interiors of proteins but do not believe in isolated ions.

Nevertheless, isolated ionized groups are found in protein interiors² (for example, Asp 102 in chymotrypsin). This fact is the central point of a paper by Quirocho *et al.* published last month in *Nature*⁶. This work identifies several highly refined cases of uncompensated charges stabilized by the surrounding protein dipoles (positive groups by the C=O dipoles and negative groups by the N-H dipoles), confirming earlier studies² emphasizing the importance of the local dipolar components.

Until recently, however, it appeared that many workers in the field attributed electrostatic effects to the macro dipoles of the protein helices (which are in fact almost completely shielded by large dielectric effects) rather than to local polarity. Apparently, despite structural information that pointed to the stabilization of isolated charges as a possible contribution to catalytic events⁷, it was not possible to examine how important such effects were (except by theoretical predictions) until the emergence of genetic engineering. Mutations of hydrogen-bond donors demonstrate that ion-dipole interactions can contribute up to 5 kcal mol⁻¹ more than the corresponding interaction in water. The reproduction of this type of electrostatic self energy by theoretical methods presents a major challenge which cannot be addressed by the simple Coulomb law-type treatments.

Early attempts to evaluate the self energies of charges in proteins were performed with semimicroscopic models that simulate the solvent by polarizable dipoles on a cubic grid². Although such models can be reduced to the continuum model (with a small grid spacing), they include the key effects of both the bound water molecules and the protein polarity, thus calculating rather than assuming the local dielectric constant. This allowed qualitative results to be obtained for self energies in protein active sites, but still did not satisfy the quantitative challenge introduced by the emergence of genetic engineering.

The availability of supercomputers,

however, changes the rules of the game. Now it is possible to use so-called 'free-energy perturbation' methods, in which the complete protein-solvent system is gradually driven through a thermodynamic cycle. Such techniques can reproduce in an almost quantitative way the changes in the free energies of ionic reaction intermediates on genetic modifications^{8,9} (see figure). The calculations are still very expensive, and it is to be hoped that refined grid models could eventually give similar results, provided that they treat the protein as a polar entity and not as an oil drop. □

Animal behaviour

Learned specializations of birds

Jared M. Diamond

MANY animal species are ecological generalists that harvest varied foods by varied foraging techniques. Does a generalist population consist of generalist individuals, all of which use the same broad repertoire of foraging techniques, or of specialist individuals, each concentrating on only a fraction of the repertoire of the whole population? A recent study by Tracey K. Werner and Thomas W. Sherry (*Proc. natn. Acad. Sci. U.S.A.* **84**, 5506-5510; 1987) provides a dramatic example of individual feeding specialization, and thereby puts the learning abilities of tropical birds into a new perspective.

Werner and Sherry studied the Cocos finch on Cocos Island, off Costa Rica in the tropical Pacific (see figure). This finch is the sole species of Darwin's finches (Geospizinae) occurring outside the Galapagos islands. As a result of the high

1. Perutz, M.F. *Science* **201**, 1187-1191 (1978).
2. Warshel, A. & Russell, S.T. *Q. Rev. Biophys.* **17**, 283-422 (1984).
3. Sternberg, M.J.E., Hayes, F.R.F., Russell, A.J., Thomas, P.J. & Fersht, A.R. *Nature* **330**, 86-88 (1987).
4. Gilson, M.K. & Honig, B.H. *Nature* **330**, 84-86 (1987).
5. Cronin, C.N., Malcolm, B.A. & Kirsh, J.F. *J. Am. chem. Soc.* **109**, 2222 (1987).
6. Quirocho, F.A., Sack, J.S. & Vyas, N.K. *Nature* **329**, 561-564 (1987).
7. Robertus, J.D. *et al. Biochemistry* **11**, 4293-4303 (1972).
8. Hwang, J.K. & Warshel, A. *Biochemistry* **26**, 2669-2673 (1987).
9. Shashidhar, N.R., Singh, U.C., Bash, P.A. & Kollman, P.A. *Nature* **328**, 551 (1987).

Arieh Warshel is Professor of Chemistry at the University of Southern California, University Park, Los Angeles, California 90089-0482, USA.

availability, variety and predictability of food resources in an aseasonal environment with very few competing species, Cocos finches have evolved extreme generalization of feeding habits. The population as a whole eats insects, crustacea, seeds, molluscs, perhaps lizards, at least 17 fruit species and nectar from at least 29 flower species, as well as insects captured by diverse techniques that include gleaning from branches or leaves or the ground, probing at branches or clusters of dead leaves, and extracting larvae from leaves.

To compare the feeding behaviour of individuals with that of the whole population, Werner and Sherry observed many colour-banded birds foraging in the same 3.4-hectare thicket during two breeding seasons and two non-breeding seasons. They assigned each feeding attempt to one of nine categories of techniques (such as probing flowers or gleaning from the ground). For 89 individuals, they classified at least 60 feeding attempts in this way, which gives a total of 26,770 feeding attempts analysed. It turns out that each individual specializes in just one or a few feeding techniques, and that 62 of the 89 individuals use just one technique for more than half their feeding attempts. For instance, 81 per cent of one male's attempts were by gleaning insects from branches, 58 per cent of another male's attempts were by probing dead leaves for crickets and cockroaches, and 42 per cent of one juvenile's attempts were by taking nectar from extrafloral nectaries. For each of the 89 individuals, the repertoire of feeding attempts is significantly more specialized than would have been observed if the same number of attempts had been drawn at random from the 26,770 attempts measured for the whole population.

Many of the usual obvious explanations for individual feeding specialization fail in

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Adult male Cocos finch probing for nectar in a *Hibiscus tiliaceus* flower. (Photograph courtesy of Tracey Werner and Tom Sherry.)

the case of Cocos finches. In this species, an individual's preferred feeding behaviour bears no relation to its sex, age or morphology, nor to the time or place of feeding. A given individual specializes in the same technique at different times of day, different months of the year and different parts of its home range, whereas different individuals feeding in the same bush specialize in different techniques. Instead, the specializations seem to be at least partly transmitted culturally, by learning from observation of other individuals. Werner and Sherry often saw juvenile finches following an adult finch at a distance of 1–2 metres and alternately watching the adult and imitating its feeding technique. Juvenile finches also imitate warblers and sandpipers. Recently fledged juveniles routinely forage in flocks from within which they can watch and imitate each other.

We take it for granted that social mammals, such as lions, acquire hunting skills only after a lengthy period of watching, imitation and practice. Birds supposedly rely on instinct and are poor at learning. But consider the implications of Werner and Sherry's work for birds living in the world's most species-rich habitats, the continental tropical rainforests. An insectivorous bird is faced with tens or hundreds of thousands of local insect species, including more than a thousand beetle species in one tree species alone (Erwin, T. *Coleopt. Bull.* **36**, 74–75; 1982). Many of these species are poisonous or spiny, others are non-poisonous but closely mimic the poisonous ones, and still others are concealed or evasive. Each species poses different problems of search and capture. Similarly, a frugivorous bird is faced with hundreds of fruit species, some of them nutritious, others poisonous and others non-poisonous but of low nutritive value. Each fruit species exhibits different external signs of when it is ripe to eat.

Although tropical entomologists and botanists devote their entire professional lives to learning to distinguish a small fraction of a local flora or insect fauna, a tropical-forest bird must master this enormous amount of local information more quickly if it is to survive. Juveniles spend months foraging with their parents after fledging, and many species forage in mixed-species flocks outside the breeding season. These behaviours could provide the schools for learning. The penalty for failing avian examinations is severe: many juvenile birds in Sarawak forest die of starvation and other causes within a few weeks after dispersing (Fogden, M. *Ibis* **114**, 307–342; 1972). Thus, tropical-rainforest birds may have been selected for learning ability. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Atlantic circulation

Surface effects on deep flow

William F. Ruddiman

To what extent is flow deep in the North Atlantic linked to climate fluctuations in the high-latitude parts of this ocean? Geochemical evidence¹ has established that variations in the Earth's orbit affect climate conditions (the Milankovitch cycles, with timescales of tens to hundreds of thousands of years). Physical oceanographers have found that variations of salinity in the deep ocean and sea surface are connected, on a timescale of decades². Now, on page 35 of this issue³, Ed Boyle and Lloyd Keigwin bridge the gap in the range of timescales, showing that flow in the deep western North Atlantic is strongly dependent on sea-surface conditions in the sub-polar Atlantic over periods of hundreds to thousands of years.

Boyle and Keigwin rely on two indicators of deep circulation measured in the carbonate shells of bottom-dwelling foraminifera obtained from sediment cores: the ratio of cadmium and calcium content (Cd/Ca), and the ratio of two isotopes of carbon, $\delta^{13}\text{C}$. Although they are in part affected by the total inventory of cadmium and of carbon-12 in the ocean, both ratios in Atlantic cores are strongly influenced by changes in deep circulation, primarily the relative rate of formation of nutrient-poor North Atlantic deep water compared with nutrient-rich deep water formed in Antarctica⁴. Deep water formed in the Atlantic today is low in Cd/Ca and rich in $\delta^{13}\text{C}$ compared with water from southern ocean sources.

In a core taken from a depth of 4,450 m on the Bermuda Rise and dated by correlation with standard oxygen-isotope ($\delta^{18}\text{O}$) stratigraphy, the rapidly deposited record reveals a distinctive signature for both geochemical indicators during the last deglaciation⁵. At the last glacial maximum, $\delta^{13}\text{C}$ values were lower and Cd/Ca values higher than today, reflecting reduced formation of nutrient-depleted North Atlantic deep water during maximum glaciation¹. Just after the beginning of the deglaciation (as marked by $\delta^{18}\text{O}$), an abrupt shift to typical interglacial values (high $\delta^{13}\text{C}$, low Cd/Ca) occurred. Then, during mid-deglaciation, there was a reversal towards glacial values (for $\delta^{13}\text{C}$, a complete reversal; for Cd/Ca, a return half way towards glacial values). Finally, late in the deglaciation, a second abrupt shift to values typical of the modern interglaciation occurred.

This geochemical signature closely resembles a sequence of North Atlantic sea-surface temperature oscillations previously detected in high-latitude sediment cores⁵. Following an interval of cold

glacial surface temperatures, the eastern and central North Atlantic abruptly warmed to full interglacial temperatures around 13,000 years ago. It then cooled to nearly full-glacial values during the 'Younger Dryas' advance of the polar front around 11,000 years ago, and finally warmed once more to full interglacial values around 10,000 years ago. Similar changes are also recorded in climatic records on continental regions next to the North Atlantic^{3,5}.

The distinctive square-wave signature in both the surface and deep-ocean signals establishes a convincing link between these two parts of the climate system on timescales of hundreds to thousands of years, although the cause of these short-term oscillations in the surface of the North Atlantic is still debated. The extent of penetration of this North Atlantic signal into the rest of the deep ocean is still to be determined. This signal may have been progressively attenuated in parts of the deep ocean with longer residence times and farther from the influence of the North Atlantic.

These results could be relevant to another problem concerning the last deglaciation. The glacial-interglacial amplitude of $\delta^{18}\text{O}$ curves measured in benthic foraminifera is largely caused by variations of global ice volume, but a quarter or more seems to arise from changes of deep-ocean temperature^{6,7}. Unfortunately, the shape of the temperature overprint during the most recent deglaciation is not known, and thus cannot be subtracted to identify the ice-volume effect. If, however, the temperature overprint has the square-wave form implied by the deep-ocean Cd/Ca and $\delta^{13}\text{C}$ data of Boyle and Keigwin³, then the stepwise decreases of $\delta^{18}\text{O}$ 14,000–13,000 and 10,000–9,000 years ago, apparent in many curves, may be artefacts of a deep-ocean warming and not indicative of sudden decreases in ice volume⁸. Similarly, the plateau in $\delta^{18}\text{O}$ values 13,000–10,000 years ago, evident to some extent in the cores studied by Boyle and Keigwin³, may be an artefact of superimposing a deep-ocean cooling event of Younger Dryas age on a background trend of gradually decreasing global ice volume.

Boyle and Keigwin also report³ important new Cd/Ca data bearing on changes in circulation at intermediate depths (1,500–2,400 m) in the North Atlantic ocean and the Caribbean. Today, the Cd/Ca ratio in the Atlantic, like the phosphorous content, varies hardly at all between 1,000 and 5,000 m depth. During the last

glaciation, the vertical Cd/Ca gradient was much larger, with higher values than today below 2,500 m and lower values above. This increased gradient seems to provide evidence of a fundamental shift away from deep-water formation and toward intermediate-water formation in the high-latitude North Atlantic.

The higher glacial Cd/Ca values below 2,500 m delineate the depth range impacted by slower rates of formation of nutrient-poor North Atlantic deep water during glaciation. The lower Cd/Ca values above 2,500 m suggest an increased rate of formation of nutrient-depleted intermediate waters typical of the Atlantic, probably because lower glacial sea-surface salinities prevented formation of the dense waters that sink to great depths today. Other geochemical evidence shows that water with different $\delta^{13}\text{C}$ characteristics existed at intermediate depths during the last glaciation, although attributed elsewhere to increased outflow of Mediterranean water^{9,10}.

Studies of deep-ocean geochemistry by Boyle, Keigwin and others in the past few

years are providing increasingly realistic three-dimensional pictures of ocean circulation during glacial climatic regimes. In the long run, this kind of careful piecing together of diagnostic evidence will bring an understanding of how changes in ocean circulation alter deep-ocean chemistry and nutrient patterns, leading ultimately to an understanding of the oceanic contribution to the carbon system, to changes in atmospheric CO_2 , and to feedbacks to global climate. □

1. Boyle, E.A. & Keigwin, L.D. *Earth planet. Sci. Lett.* **76**, 135–150 (1987).
2. Swift, J.H. in *Am. geophys. Un. Monogr.* **29**, 39–47 (1984).
3. Boyle, E.A. & Keigwin, L.D. *Nature* **330**, 35–40 (1987).
4. Boyle, E.A. *Geochim. cosmochim. Acta* **50**, 265–276 (1986).
5. Ruddiman, W.F., Sancetta, C.D. & McIntyre, A. *Phil. Trans. R. Soc. B280*, 119–142 (1977).
6. Shackleton, N.J. & Opdyke, N.D. *Quat. Res.* **34**, 39–55 (1973).
7. Chappell, J. & Shackleton, N.J. *Nature* **324**, 137–138 (1986).
8. Ruddiman, W.F. in *North America and Adjacent Oceans during the Last Deglaciation* 463–478 (Geol. Soc. Am., 1987).
9. Zahn, R., Winn, K. & Sarinthein, M. *J. Mar. Res.* **41**, 327–347 (1983).
10. Oppo, D. & Fairbanks, R.G. *Earth planet. Sci. Lett.* (in the press).

William F. Ruddiman is at the Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA.

Plant ecology

Growth and foraging behaviour

William J. Sutherland

In most closed plant communities, establishment of seedlings is rare; much of the spatial change in the vegetation is caused by clonal growth, for example, by the production of stolons and rhizomes that can vary in both length and persistence. By clonal growth, an individual plant can disintegrate and disperse. The growing plant encounters a mosaic of environments caused by cowpats, mole hills, competitors, soil heterogeneity and many other sources of disturbance. Slade and Hutchings have recently examined^{1–3} the exact manner by which the growth pattern of a stoloniferous herb responds to patchiness, and draw analogies between the foraging behaviour of searching animals that respond to patches by modifying their search path, and the growth pattern of clonally growing plants that respond to patches by modifying their morphology.

Zoologists find it useful to consider the selection pressures acting on morphological and behavioural traits by assessing the costs and benefits of the possible solutions and predicting which particular trade-off between costs and benefits will give the maximum net benefit to the individual. Large pores in eggshells, for example, lose more water but allow more oxygen to diffuse in. Alexander⁴ calculated the optimal compromise and showed that the relationship between pore size and egg size (for species ranging from the wren to the ostrich) can be predicted.

Botanists have recently been exploring the costs and benefits associated with various plant traits to understand plant morphology; Givinish⁵, for example, has attempted to predict the branching angle and branching height of the sexual shoots of a forest herb *Podophyllum peltatum*. An increase in branching angle has two conflicting consequences: vein costs decrease because the two leaves can become more symmetrical without overlapping; whereas stem costs increase because the length of the branches in each stem increases. Givinish produced a quantitative model to calculate the branching angle that minimizes the total support costs, and then used a similar approach, invoking the pipeline theory of economics, to determine the branching height that minimizes the total support costs. In both cases, the theory provides a very satisfactory prediction of the morphology. Further analyses explain why the sexual shoots branch and why they bear two leaves rather than one larger leaf. Such economic analyses can be used⁶ to elucidate the selective pressures acting on characteristics such as leaf reflectivity, effective leaf size, stomatal conductance, size of photosynthetic enzyme pools, crown form, xylem structure and allocation to roots versus shoots.

Analysis of the foraging behaviour of animals by assessing the consequences of various decisions is one successful application of economic models⁷. Several people

have suggested that a similar approach could be applied to the study of plant growth and that the growth pattern of a stoloniferous or rhizomatous species is analogous to the search path of a foraging animal^{8,9}. Just as modification of its search path enables an animal to concentrate its time in the most lucrative patches, modification of the growth pattern may result in plant ramets accumulating in the better sites.

Plants with stoloniferous growth may respond to environmental heterogeneity such that most ramets accumulate within the favourable sites. Salzman¹⁰ placed ramets of *Ambrosia psilostachya* in a saline gradient and showed that as a consequence of the subsequent growth most of them accumulate at the less saline end. Eriksson¹¹ showed that *Potentilla anserina* predominates in the temporary areas of open ground and that the species preferentially colonizes areas cleared of competitors. Turkington and Harper¹² found that clones of clover track the continually changing distributions of the various grass species. These data demonstrate that redistribution occurs by vegetative growth, but they do not show the precise manner by which this redistribution takes place.

Slade and Hutchings^{1–3} now show the exact manner in which morphology is modified by the immediate environment of a woodland herbaceous plant *Glechoma hederacea*, which proliferates by producing pairs of secondary stolons from the two lateral meristem buds at the stolon nodes. The authors rooted individual ramets, allocated a nutrient and light regime, and recorded the subsequent pattern of growth. The crucial variables turn out to be the branching probability, internode length and branching angle. When few nutrients are supplied, the stolons produce fewer branches but longer internodes¹. Light has comparable effects on growth, with fewer branches and longer internodes being produced when photon-flux density is low². Slade and Hutchings interpret these and other data using foraging theory. They suggest that nutrient-rich sites are exploited by a growth form that consolidates the clone's occupation of the area, whereas the response to poor surroundings will maximize the probability of the clone growing and expanding elsewhere.

The morphology of stoloniferous plants also responds to the density of competitors. When white clover, *Trifolium repens*, is grown with grasses, the ramets possess fewer stolon branches and longer internodes than when grown alone and, as a consequence, growth is primarily linear¹³. When *Hieracium pilosella* is grown at a range of densities, the number of primary stolons and the extent of stolon branching are strongly density dependent¹⁴.

What is the physiological basis of foraging? The morphological responses of

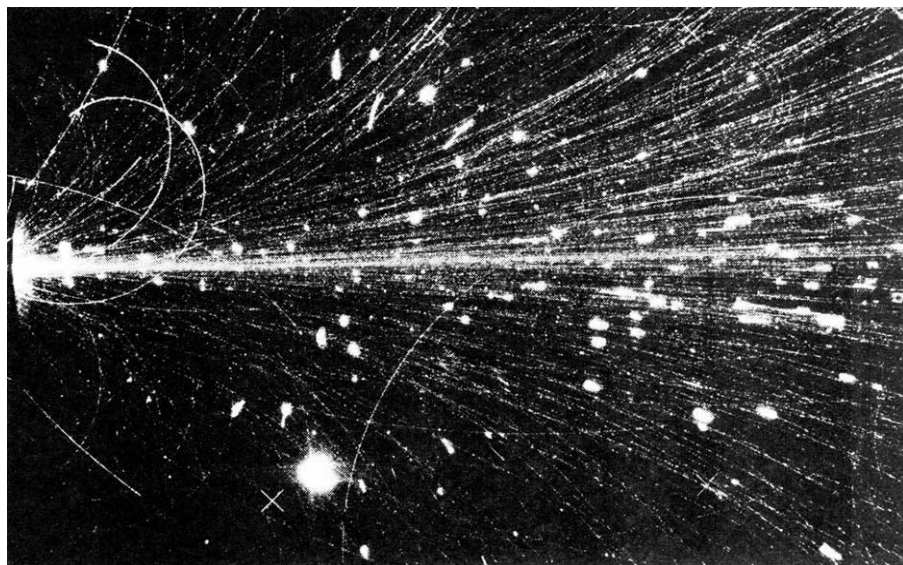
plants to light, nutrients and competition may be easier to interpret than the equivalent behavioural responses of animals. The light underneath a plant canopy has a high infrared/red ratio because of differential absorption by the leaves. A high ratio of light results in increased apical dominance, inhibition of the lateral buds and elongation of the internodes⁶. To see whether the change in spectra caused by real canopies are important, Solongarachi and Harper¹³ grew *T. repens* under canopies of clover leaves floating on shallow tanks of water and compared the growth with that of plants under canopies of black polythene 'leaves'. The polythene 'leaves' reduced the intensity of photosynthetically active radiation without altering the infrared/red ratio. Both the quantity and quality of light have a considerable effect on the branching such that it is almost wholly suppressed under the dense clover canopy.

Although the studies of *G. hederacea* are the most comprehensive to date and provide data that are readily compatible with the ideas derived from animal foraging theory, it would be unwise to assume that all rhizomatous and stoloniferous species behave in the same manner. Indeed, the observation that internodes are longer in nutrient-poor sites is contradicted by many other studies^{11,16-18}. It could be that stolons must extend outside the resource depletion zone¹⁹. In plants with determinate growth, such as *G. hederacea*, this zone may vary little whereas species with indeterminate growth may have to produce longer stolons to evade the larger resource-depletion zone resulting from the enhanced growth under the better conditions. It is clear that understanding the distribution, dynamics and abundance of stoloniferous and rhizomatous plants requires detailed work on how morphology is modified by environment. □

- Slade, A.J. & Hutchings, M.J. *J. Ecol.* **75**, 95–112 (1987).
- Slade, A.J. & Hutchings, M.J. *J. Ecol.* (in the press).
- Hutchings, M.J. & Slade, A.J. in *Plant Population Biology* (eds Davy, A.J. *et al.*) (Blackwell, in the press).
- Alexander, R. *McN. Optima for Animals* (Arnold, London, 1982).
- Givnish, T.J. in *On the Economy of Plant Form and Function* (ed. Givnish, T.J.) 525–583 (Cambridge University Press, 1986).
- Givnish, T.J. (ed.) *On the Economy of Plant Form and Function* (Cambridge University Press, 1986).
- Stephens, D.W. & Krebs, J.R. *Foraging Theory* (Princeton University Press, 1986).
- Bell, A.D. *Perspectives in Plant Population Ecology* (ed. Dirzo, R. & Sarukhan, J.) 48–65 (Sinauer, Mass. 1984).
- Grime, J.P. *Plant Strategies and Vegetation Processes* (Wiley, Chichester, 1979).
- Salzman, A.G. *Science* **228**, 603–604 (1985).
- Eriksson, O. *Oikos* **46**, 82–87 (1986).
- Turkington, R. & Harper, J.L. *J. Ecol.* **67**, 201–218 (1979).
- Solongarachi, S.M. & Harper, J.L. *Oecologia* **72**, 372–376 (1987).
- Bishop, G.F. & Davy, A.J. *Oecologia* **66**, 417–422 (1985).
- Smith, H.A. *Rev. Plant Physiol.* **33**, 481–518 (1982).
- Lovett Doust, L. *J. Ecol.* **75**, 555–568 (1987).
- Ginzo, H.D. & Lovell, P.H. *Ann. Bot.* **37**, 753–764 (1973).
- Sackville Hamilton, N.R. thesis, Univ. Cambridge (1980).
- Harper, J.L. in *Population Biology and Evolution of Clonal Organisms* (eds Jackson, J.B.C., Buss, L.W. & Cook, R.E.) 1–33 (Yale University Press, New Haven, 1985).

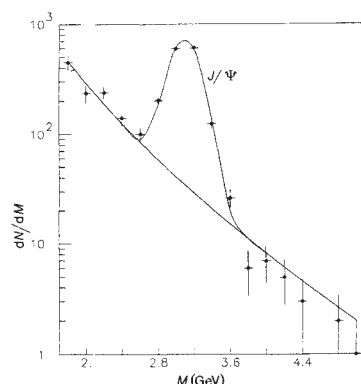
William J. Sutherland is in the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Signature of quark–gluon plasma seen



THE picture above shows the debris thrown out when ultra-relativistic sulphur ions smash into stationary gold nuclei (photograph courtesy of R. Stock). It was recently obtained at CERN in Geneva by a collaboration using a 'streamer' chamber that detects all charged particles produced in such collisions (up to 600). The group is one of six that are trying to recreate the conditions of the Universe just seconds after the Big Bang, when matter was so hot and dense that the quarks and gluons that make up nuclear matter were 'unconfined'.

The experiments involve accelerating oxygen ions (last autumn) and sulphur ions (last month) to an enormous energy of 200 GeV per nucleon in CERN's Super Proton Synchrotron (SPS). This year's results have yet to be analysed. Results from last year's oxygen experiments, presented at the recent *Quarkmatter 87* conference, are



consistent with the creation of a quark–gluon plasma, although they are not conclusive. The graph (courtesy of P. Sonderegger) shows the rate of production dN/dM of muon pairs ($\mu^+ + \mu^-$) as a function of their total mass M . The resonance peak arises from pairs produced by the decay of J/ψ mesons. As described recently

in a News and Views article by Helmut Satz (*Nature* **327**, 660; 1987), J/ψ production should be suppressed in a quark–gluon plasma, and such a suppression (up to 50 per cent) is seen in oxygen–uranium collisions, especially those selected for 'directness'. Physicists are wary, however, in case another, unknown mechanism is at work.

Streamer chambers should show another signature of quark–gluon plasmas — a change in the relative production rates of Λ^- and π^- mesons. These data are still to be analysed. But by correlating all possible pairs of pions in each picture, physicists at CERN find that the 'fireball' created in the collision is almost spherical (slightly flattened along the collision axis) and lasts about 6.4×10^{-23} s. Both conditions are necessary to form a quark–gluon plasma — the latter so that the collision energy has time to spread throughout the fireball. This pion interferometry (which can be likened to astronomical Hanbury-Brown and Twiss interferometry) shows that the fireball expands to about $6.5-8 \times 10^{-15}$ m before it 'freezes out' into individual hadrons made of confined quarks.

With larger nuclei, or higher energies, the multiplicity of pions produced should be large enough for each collision to be analysed individually with reliable statistics, and the results compared. The average direct collision of oxygen on gold involves all 16 nucleons of the projectile oxygen and 50 of the target gold. Using sulphur ions, this increases to $(32+80)$ nucleons. The energy deposited is about 20 GeV per nucleon pair, and is about 1.8 times more with sulphur than with oxygen. At a proposed ion collider at Brookhaven National Laboratory, United States (in which both ions are moving), the usable energy would be increased about tenfold. Scientists at CERN are proposing a scheme to accelerate much heavier lead ions in the SPS to create the hotter, more dense conditions. □

Atomic physics

Electron-scattering experiments

H. Kleinpoppen

ELECTRON-EXCHANGE and spin-orbit effects are inherent in the quantum description of atoms. In 1926 Heisenberg explained the existence of two species of helium, ortho- and para-helium, by considering exchange interactions between the two electrons of the atom and also spin-orbit effects. Such spin effects also appear in electron-atom scattering, and these have been studied in various ways since the early 1970s. An interesting result is now reported by J.J. McClelland, M.H. Kelly and R.J. Celotta (*Phys. Rev. Lett.* **58**, 2198–2200; 1987), who have measured for the first time the competition between exchange and spin-orbit interactions in collisions of electrons with sodium atoms.

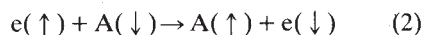
It is spin-orbit interactions that account for the splitting of the helium atomic-energy levels. The two electrons can have their spins parallel (spin-triplet or $S=1$ states) or antiparallel (spin-singlet or $S=0$ states). The magnetic moment of the spinning electrons interacts with the orbital motion as $\mathbf{L} \cdot \mathbf{S}$, where $\mathbf{L}$ is the orbital angular momentum.

For atomic collisions, electron-spin effects are also important in different types of energetic interactions such as Coulomb-direct, Coulomb-exchange and spin-orbit. They can be investigated by using polarized electrons and polarized atoms as initial scattering particles and measuring of the spin of the electrons and/or the atom after the collision, and also by measuring a collisional asymmetry effect of the scattered electron. In their recent paper McClelland *et al.* report a study of exchange and spin-orbit effects in collisions between partially polarized sodium atoms and electrons (Fig. 1).

The scattering of spin-polarized electrons by spin-polarized 'one-electron' atoms (such as hydrogen or the alkali atoms) can be described as follows. The spin polarization of the electrons and atoms is denoted by specifying the direction of the spin orientation as parallel ($\uparrow$) or antiparallel ($\downarrow$) to an arbitrary axis of quantization (more loosely described as spin-up or spin-down). Typical scattering reactions can be written as:

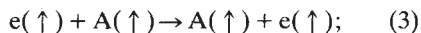


or



where A represents the atom. Reaction (1) is the Coulomb-direct process in which there is no electron or spin exchange. Reaction (2) is the Coulomb-exchange process in which the two electrons exchange each other without changing

their spin directions. A third scattering reaction, with the spins initially parallel to each other, would not allow changes in the spin directions under a spin conservation law:



Quantum mechanically these three reactions can be described by the complex scattering amplitudes f, g and $f-g$, respectively, giving all possible information on the process. The squared amplitude of a given interaction is proportional to the scattered intensity of the electrons. Alternatively, the electron-atom collision process can be described by singlet (S) and triplet (T) scattering amplitudes depending on the joint antiparallel or parallel spin directions of the collisional and target electron, respectively, with $f = \frac{1}{2}(S+T)$, $g = \frac{1}{2}(S-T)$.

In their experiment, McClelland *et al.* measured the intensities of the scattered electrons as a function of the electrons' initial energy and scattering angle. They did so for the four possible combinations of spin orientation: spins parallel, up or down; and the spins antiparallel (again with two possibilities, each one up or down). They define an asymmetry factor A^{exch} for the exchange process as the difference between the intensities for the parallel and antiparallel processes (normalized to the sum of the intensities, and the degree of polarization of the initial beams).

This measured quantity can be related to the scattering amplitudes of interest by

$$A^{\text{exch}} \propto \frac{\text{Re}(f^*g)}{\sigma} = \frac{1 - |T|^2/|S|^2}{1 + 3|T|^2/|S|^2} \quad (4)$$

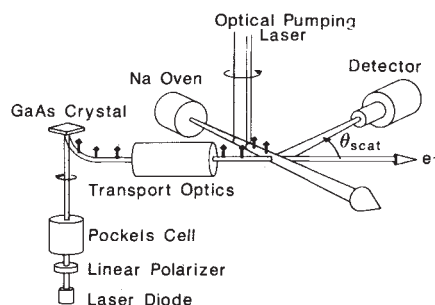


Fig. 1 Apparatus in the electron-atom scattering of McClelland *et al.* Photoelectrons are generated from the GaAs crystal by circularly polarized laser light. (The Pockels cell alternates the direction of polarization.) Similarly, the sodium atoms are polarized by 'optical pumping' using a second laser. In both cases, the polarization of the laser light is efficiently transferred to the absorber. The scattered electrons are collected by a movable detector.

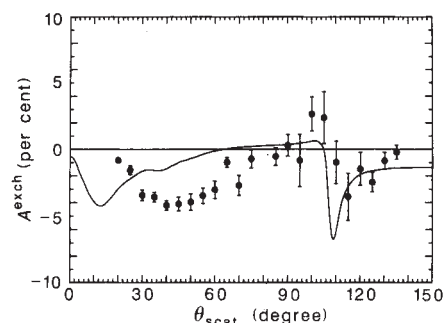


Fig. 2 Exchange asymmetry A^{exch} as a function of electron scattering angle. Experimental data points by McClelland *et al.*; theoretical curve represents a two-state close coupling calculation. The error bars are one standard deviation derived from counting statistics.

where σ is the differential cross-section, a function of the kinetic energy of the electrons and their scattering angle.

Figure 2 shows measured data of this exchange asymmetry A^{exch} for sodium atoms at incident electron energy of 54.4 eV as a function of the electron scattering angle θ_{scat} . The negative values of A^{exch} between 30 and 60° can be attributed to a negative term of interference between the exchange- and direct-scattering amplitude or alternatively as a slight dominance of the triplet-scattering amplitude. The 'oscillation' around 108° can be interpreted as the result of the singlet and triplet cross sections having minima at slightly different scattering angles.

Considerable discrepancies are seen, however, between theory and experiment for A^{exch} , although their behaviour is qualitatively similar. Are the discrepancies the result of inadequate treatment of exchange interactions in the calculation or of the combined effects of exchange and spin-orbit interactions? Spin-orbit effects result from the orbital motion of the projectile electron through an electric field (caused by the distortion of the atoms) and the interaction of the magnetic moment of the electron with the motional magnetic field.

Substantial influences from spin-orbit interactions on the electron-scattering process have been reported with heavier closed-shell atoms as targets where electron-spin experiments lead to a complete description of the collision process (for example, Wübker, W., Möllenkamp, R. & Kessler, J. *Phys. Rev. Lett.* **49**, 272; 1983). Such spin-orbit interactions are particularly important in electron scattering with heavy atoms but they have generally been ignored in experiments using lighter atoms. The new experiment of McClelland *et al.*, especially the data close to $\theta_{\text{scat}} = 108^\circ$, shows that asymmetry measurements can allow a complete description of electron collisions with light atoms. □

H. Kleinpoppen is in the Atomic Physics Laboratory, University of Stirling, Stirling FK9 4LA, UK.

A superconductivity primer

Philip Campbell

High-temperature superconductors continue to tantalize the experts. Here is an introduction for others to both conventional and novel varieties.

SINCE the report a year or so ago, by J.G. Bednorz and K.A. Müller, that certain transition metal oxides are superconductors at unprecedentedly high temperatures, there has been a torrent of experimental and theoretical research aimed at defining the characteristics of these materials, understanding the mechanisms involved, and finding materials with transition temperatures even greater than the 95 K or so at which experimentalists now appear to be stuck. It was quickly realized that the theory of conventional superconductivity cannot account for the properties of the new materials. Whether an extension of present theories or a significantly new approach is necessary remains to be seen.

One remarkable feature of the past year has been the extent to which people from a variety of disciplines have been drawn into research in superconductivity. For that reason, not to mention the great interest that has been aroused among non-specialists, there may be a need for a summary in plain language of the basic attributes of conventional superconductors and the mechanisms involved, as well as the distinguishing features of the new materials.

What follows is not a formal review — no references are provided. The hope is simply that readers of this account may be better able to follow *Nature's* past and future coverage of high-temperature superconductors and also next week's Boston meeting on the topic. A subsequent article will dwell on the distinctive characteristics of the new materials and the range of explanations being considered.

Electrical resistance

The most striking attribute of a superconductor is, of course, that its resistance to a direct current is zero, as far as can be measured. In one experiment, a current established in a superconducting ring was shown to have a decay time of not less than 100,000 years.

Seventy-five or so years after the discovery of superconductivity this 'perpetual motion' still seems hard to credit. After all, in a conventional conductor carrying a steady current, there is an equilibrium established between the force on the current-carrying electrons from the electric field and the resistance that they encounter arising from interactions with the atomic lattice. As it happens, a perfectly periodic and rigid lattice would

provide no resistance at all because, in general, the scattering effects of neighbouring ions would destructively interfere and cancel. Resistance arises principally because thermal vibrations of the lattice and the presence of defects or impurities distort the regularity and give rise to net scattering.

Copper is a good example. Very pure copper crystals at liquid-helium temperature (4 K) have a conductivity nearly 100,000 times that measured at room temperature. The quantized lattice vibrations ('phonons') accounting for the enhanced resistance at higher temperatures individually scatter electrons and, indeed, can exchange energy and momentum with them. The stronger the electron-phonon interactions, the poorer the conductor.

Some of the electrical and thermal properties of ordinary conductors are well represented by the 'free electron' model, in which there is a sea of conduction electrons unaffected by the electric charges of the ionic lattice and scattered only rarely by each other. There is an infinity of states accessible to electrons, which are essentially those of the textbook quantum mechanics problem of an 'electron in a box'. At absolute zero (that is, in the ground state) the conduction electrons fill the lowest available energy levels up to a threshold known as the Fermi energy, E_F . As the temperature and thus the kinetic energy of the conduction electrons increase, higher levels become occupied at the expense of less energetic states.

The Fermi energy of metals is typically a hundred times larger than the average random thermal energy (kT , where k is Boltzmann's constant and T the temperature), so that only a small proportion of the electrons is raised in energy in this way, from just below to just above the Fermi energy level.

The free-electron theory provides a qualitative picture of conduction in the metallic state. In the ground state, the wave vectors of the electrons (solutions of the free-electron Schrödinger equation, encapsulating the electron distribution and behaviour) define a sphere in vector (momentum) space whose surface (defined by the maximum vector amplitude and corresponding to the Fermi energy) is called the Fermi surface.

The velocity of electrons at the Fermi surface is an important parameter: multiplied by the characteristic collision time, it gives the mean free electron path, which

for copper is three millionths of a centimetre at room temperature and 0.3 cm at 4 K. Another important quantity is the number of different energy states per unit energy interval (the 'density of states') at the Fermi energy, which among other things is linked with the contribution by electrons to the specific heat of the metal.

While the free-electron model can account for many aspects of metallic behaviour, it does not explain why some materials are insulators and why others, the semiconductors, will conduct only when electrons acquire a minimum additional energy (as, for example, when the temperature is increased). But elaborations of this model which allow for the motion of the electrons in the periodic electric field of the lattice ions (the 'nearly-free-electron model'), show that in some crystal lattices, electrons of particular energies cannot propagate. This phenomenon underlies the existence of semiconductor 'band gaps', for example.

Electron-phonon interactions contribute to electrical resistance directly, but there are other ways in which the interaction between electrons and the lattice may affect the electrical properties of a metal. In particular, the mass of an electron may be effectively increased, both in metals and in insulators.

For example, in an ionic solid such as potassium chloride, an electron will interact repulsively with the chlorine anions and attractively with the potassium cations, thereby setting up a strain in the lattice. The combination of a local region of strain and the electron is referred to as a 'polaron'. The effect of these interactions is to increase the inertia and thus the effective mass of the electron. If the lattice strain is small (the case of the 'large polaron') the enhancement of mass is small and the electron moves freely within a conduction band. But the strain may be large ('small polaron'), in which case the electron will spend most of its time trapped by ions, tunnelling only slowly (or, at higher temperatures, hopping) through the lattice.

Superconducting metals

Two criteria must be satisfied if a material is to be classed as a superconductor: the electrical resistance to a direct current must be literally zero and the material must expel any external magnetic flux as it is cooled through its transition temperature in a magnetic field — the Meissner

effect. The former of these two properties was that first recognized when, in 1911, H. Kamerlingh-Onnes examined the low-temperature properties of mercury. It took another 40 years or so for the effect to be fully explained.

Even before there was a satisfactory microscopic theory of superconductivity, it was clear that not all electrons in a superconductor are superconducting. The fraction of electrons conducting in a normal fashion increases gradually from zero at absolute zero to unity at the transition temperature. This explains why a superconductor, while exhibiting zero direct current (d.c.) resistance, does have a finite impedance to alternating currents.

The superconducting electrons which 'short circuit' a static electrical field must by some means be immune from the scattering by impurities and lattice vibrations which resist the movement of the ordinary electrons. A crucial clue to this mechanism was the discovery of the 'isotope effect': mercury specimens with different isotopic composition were found to have different superconducting transition temperatures varying (inversely) with the atomic mass, indicating that an electron-phonon interaction must be involved.

As it turns out, the infinite conductivity can be (and was historically) explained in two steps: (1) by demonstrating that it is a natural consequence of the 'superconductor energy gap' — a gap between the ground-state energy levels of the superconducting electrons and the minimum level to which they can be excited, and (2) by explaining the existence of such a gap in superconducting materials, as was finally achieved by J. Bardeen, L.N. Cooper and J.R. Schrieffer (BCS) in 1957.

For general understanding, the two steps are best inverted. BCS showed that, through the mediation of the electron-phonon interactions, there can be cooperative states of the system of all the electrons in a superconducting sample.

The cooperative states, being states of all the electrons, are not easily described except in terms of the simplified analysis due to L.N. Cooper (which preceded the full elaboration of BCS). First, electrons interact with the lattice which contains them through the Coulomb attraction of the ions, and thus also interact with phonons, the vibrational quanta of the lattice.

This process leads to an interaction between pairs of electrons which is distinct from but analogous to the electrostatic repulsion between them. In the language of quantum field theory, just as the electromagnetic interaction between two electrons is represented by the emission of a 'virtual' phonon by one and its absorption by the other, so the phonon interaction is represented by the emission and absorption of lattice phonons.

What are the forces that result? The

calculations show that in suitable circumstances, the phonon-mediated forces between pairs of electrons can be attractive. One condition that must be satisfied for this to be the case is that the components of an electron pair must have opposite spins and equal and opposite momenta. The pair will be 'stable' so long as the electron thermal energy is less than the pair binding energy — a criterion that leads to an approximate determination of the critical temperature, at which these quantities are equal.

Cooper found that the spatial extent of such electron pairs might be of the order of 10^{-6} cm, much greater than the inter-ionic separation in the lattice. The strength of the coupling is determined by that of the electron-phonon interaction. This quantity determines electrical resistance in a normal metal, which explains why metallic superconductors are poor conductors above the superconducting transition temperature.

Plainly, the average energy (potential plus kinetic) of the superconducting electrons in a metal must be less than E_F — otherwise the cooperative state would not be formed. It turns out that the extra kinetic energy of the superconducting electrons is more than offset by their mutual binding energy due to electron-phonon coupling. BCS generalized Cooper's analysis for the 10^{23} 'free' electrons per cc that are found in a typical metal and thereby successfully demonstrated that such collective states can exist. For example, in a superconductor at absolute zero, when all the electrons are superconducting, they may be thought of as occupying the energy levels of a separate superconducting electron band with a maximum energy known as the 'superconductive ground-state energy', or E_0 , for short, which lies below E_F .

BCS shows that superconducting electrons can be excited above E_0 , but that the excited energy levels cannot fall below a certain minimum, which as it happens lies above the Fermi energy of the conventional metal. The gap separating the ground-state energy and this minimum is the 'energy gap' of a superconductor. (It should not be confused with the forbidden energy regions of semiconductors.)

As the temperature increases, this gap decreases, gradually at first and then rapidly to zero at the transition temperature. The gap manifests itself most obviously through the exponentially increasing specific heat of the material up to the discontinuity at the transition temperature as well as in the optical properties of superconductors.

One specific prediction of BCS is that the ground-state energy gap should be about 3.1 times the electron thermal energy at the transition temperature. The measurements of this ratio for high-temperature superconductors have proved

troublesome, but all the signs are that it is significantly greater. The logic of BCS would imply that the pairing mechanism must involve stronger coupling than occurs in superconducting metals.

What happens if a steady current is established in a loop of superconductor in its ground state? The maximum energy level is now slightly increased because of the extra kinetic energy, but the entire population of superconducting electrons still fills the available states at or below that threshold. For one of those electrons to achieve a different state after scattering by a phonon, for example, it must bridge the energy gap. But there is insufficient energy available, so that the scattering cannot occur, with the consequence that the material offers no electrical resistance.

Two aspects of the BCS explanation of metallic superconductivity highlight the difficulties of visualizing the mechanisms involved. First, the electrons cannot be described as distinct point-like entities in space but, as described by the quantum mechanical formalism of BCS, are 'delocalized'. Second, it is the momentum characteristics of electrons as mentioned above, not their physical location, that determines pair formation. For this reason, the BCS interactions can be described as 'momentum-space' pairing.

In contrast, some of the pairing mechanisms suggested for the new oxide superconductors arise in 'real space'. Be that as it may, it turns out that the collective behaviour of an ensemble of paired electrons can most conveniently be described by a single many-particle 'macroscopic' wave-function. As will be seen, it is this aspect of superconducting electron behaviour that readily accounts for some of the most striking characteristics of superconducting devices.

The BCS treatment thus provides an explanation of zero resistance. It also quantifies the maximum current, in principle, that a superconductor can carry without losing its superconducting properties: the additional kinetic energy of electrons carrying a current can reach the point at which it exceeds that required to break the 'Cooper' pairs. In practice the critical currents of superconductors are smaller, sometimes by many orders of magnitude, than this limit because their magnetic side-effects are also detrimental to superconductivity.

Subsequent work has shown that, for all its power, BCS cannot explain some of the fine detail of metallic superconducting behaviour because it assumes instantaneous pairings. But the phonons generated by an individual electron travel maybe hundreds of times more slowly than the 10^6 m s⁻¹ (the 'Fermi velocity') of typical electrons. The pairing interaction (first quantified in its time-retarded form by Eliashberg) is thus sometimes depicted in terms of messages left by electrons for one

another in the form of lattice polarization.

It is too early to tell, but one of the crucial differences between the metallic and the new oxide superconductors may be that the pairing mechanism for the latter is effectively instantaneous. Certainly, most theorists have concluded that the BCS mechanism can, at most, represent only a small contributing factor at superconducting temperatures above 40 K. But several unique aspects of these materials — their anisotropic structure and the apparently crucial influence of oxidation state are but two — severely complicate the lives of those attempting to elucidate the underlying processes.

Magnetic behaviour

The response of a superconductor to external magnetic fields most vividly illustrates how far removed from classical electromagnetism are the underlying processes. For a 'conventional' conductor, the bulk magnetic behaviour can be predicted using Maxwell's equations, with the additional constraints provided by Ohm's law, to show that, as a sample undergoes a transition to zero resistance, the internal magnetic field will be trapped.

The fact that a superconductor, by contrast, expels the field implies that the standard picture of a sea of conduction electrons is inappropriate. The underlying reason for the characteristic behaviour of superconductors is that the wave-functions of the superconducting states in effect describe ensembles of pairs of electrons with opposite spins which, as entities, are not restricted by the Pauli exclusion principle that no two may be in exactly the same state. The consequence is that the superconducting electrons within a sample in a static magnetic field will respond by forming a pattern of surface currents which actually cancels out the external field. That this would happen was surmised long before BCS by F. and H. London in 1935.

This behaviour has been called 'perfect diamagnetism', but it differs from conventional diamagnetism in that macroscopic surface currents, rather than changes of atomic magnetic moments, are its driving force. The Londons' analysis showed that the field decays exponentially into the sample over a characteristic 'penetration depth', λ , typically for metals between 10 and 100 nm.

The Londons' simple description was found by A.B. Pippard to be insufficient to account for the way in which the penetration length for a pure superconducting metal (tin) depends on the orientation of the surface relative to the (tetragonal) crystal axes. He showed that the current flowing at any particular point would be a function of the magnetic field within a certain distance of that point, called the 'coherence length'. The London equations, in contrast, relate the current

to the magnetic field only at the point itself. (More precisely, the London and Pippard equations relate the current density to the magnetic vector potential.) As will be seen, the coherence length is a crucial characteristic of any superconducting material.

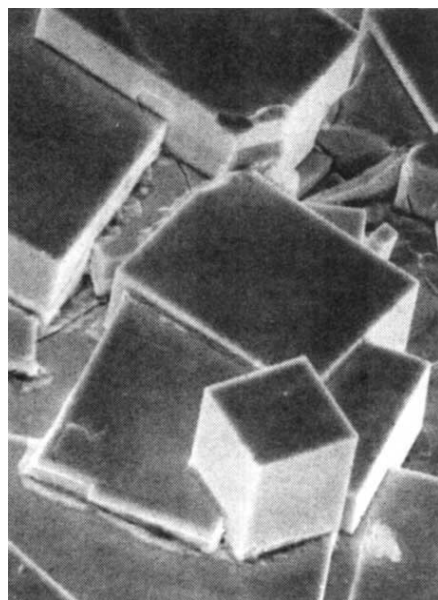
There are two types of superconductor, distinguished by their response to an applied magnetic field. Type I superconductors exhibit perfect diamagnetism up to a critical applied field, at which point the superconductivity is lost and the average magnetization of a sample rises abruptly from -1 to the value (near zero) of the normal material. (The abruptness depends to some extent on the shape of the sample, the ideal transition occurring for a thin rod or wire.) In type II materials the magnetization is governed by two critical values of the applied field; from -1 at the lower threshold it gradually increases, only reaching the normal value at the upper critical field, which in some cases is one hundred times stronger than the lower. Between the two thresholds (the 'mixed' state), magnetic flux penetrates the sample in discrete threads and superconductivity persists.

Typical values of critical fields for type I superconductors (for example pure metals) lie roughly between 0.01 and 0.1 tesla. The value quoted applies to absolute zero, the critical field decreasing to zero as the temperature is raised to the transition value. Type II superconductors (which, apart from vanadium and niobium, are alloys and compounds) have lower critical field intensities similar to the type I values above, whereas the upper critical field strength may amount to several teslas or more.

The new oxide superconductors are type II materials with upper critical fields typically greater than 100 teslas. Their anisotropic structures combined with the difficulties of producing superconducting single crystals have so far rendered it impossible to achieve a consensus as to values of penetration depths or coherence lengths.

Phase transitions

As far as conventional superconductors are concerned, the BCS theory can explain all these phenomena and more, but considerable insight was obtained in advance of that by Ginzburg and Landau who, in 1950, analysed the energetics of the superconducting phase changes. Landau had previously produced a general theory of second-order phase transitions (in which the entire system changes state essentially simultaneously without the involvement of latent heat). The essential concept of Landau's theory is the 'order parameter', having its maximum magnitude when the system is most ordered (a lattice of parallel atomic dipoles in a magnet for example), as might



Crystals of superconducting $\text{YBa}_2\text{Cu}_3\text{O}_7$. Most samples of high-temperature superconductors have been polycrystalline, and the resulting presence of grain boundaries has hampered investigations (see text). Fractionation during solidification of molten oxides hinders the growth of large single crystals of superconductors. The crystal shapes seen above reflect the orthorhombic crystal structure.

occur at absolute zero. As temperature increases, the order parameter decreases, becoming zero (maximum disorder) at and above the critical temperature. To explain phenomena in the neighbourhood of the transition, Landau formalized the free-energy density of the system in terms of the order parameter and thus determined the value of the parameter associated with the minimum free energy at a given temperature.

Ginzburg and Landau applied this theory to superconductors. The condition for a minimum of free energy, it turns out, closely resembles a Schrödinger equation, thus allowing the identification of the order parameter with the macroscopic wave-function mentioned above. The solutions reveal several features of superconductors: they reproduce the London/Pippard equations; they yield a characteristic minimum length over which the wave-function can significantly change — the coherence length; and they show that the wave-function will extend out of the superconductor at a boundary — in other words, electron pairs can tunnel out of the superconductor in certain circumstances. This phenomenon is crucial to the Josephson effect (see below) which underlies so many applications of superconductors.

Perhaps the most significant insight provided by the Ginzburg-Landau theory concerns the difference between type I and type II superconductors as exhibited in their contrasting responses to external magnetic fields. The expulsion of magnetic flux by a sample costs energy, and there will clearly be an applied field

strength at which it becomes energetically easier to assume the normal conducting state.

Loosely speaking, Ginzburg and Landau showed that if — as in type II materials — the penetration depth exceeds the coherence length, the increase in energy of the bulk from the formation of restricted regions of normal conducting material is more than offset by the decrease in magnetic energy as the flux enters those regions. Thus normal 'cores' are formed, lying parallel to the applied field, sheathed by vortices of supercurrent and threaded by magnetic flux.

It turns out that a triangular lattice of cores is the preferred geometry in this mixed state. At the lower critical field the normal cores are spaced a penetration depth apart. As the field increases, more cores are formed and the separation decreases until, at the upper critical field, the cores are separated only by a coherence length or so.

This 'vortex motion' can be inhibited if there are impurities or defects such as grain boundaries present. As the field increases above the lower critical threshold, such 'flux pinning' (which is yet to be fully explained at the microscopic level) inhibits the inward motion of newly formed cores, thus generating a magnetic-field gradient within the material boundary which must, by Maxwell's equations, be associated with an enhanced current flow. Thus control of microstructure can allow one to engineer materials of comparatively high critical currents, with clear practical benefits.

The inhibition of homogenous flux penetration by pinning renders the upper critical field transition irreversible — thence the terminology 'reversible' and 'irreversible' type II materials.

Flux quantization

The macroscopic wave-function that describes the superconducting state has a clearly defined phase and amplitude. An analysis of its behaviour in a superconducting loop shows that the magnetic flux that can thread the loop is quantized. The quantum units ('fluxoids') have a value $h/(2e) \approx 2.10^{-15} \text{ T m}^2$, where h is Planck's constant and e is the electronic charge. The appearance of $2e$ as the denominator reflects the electron pairing involved. It was a flux quantization experiment of this sort that confirmed that some type of electron pairing underlies the behaviour of the new high-temperature superconductors.

Josephson effects

If two superconductors are separated by a sufficiently thin layer of insulator (two metal layers separated by an oxide, for example), the penetration of the barrier by the macroscopic wave-functions leads to overlap, or electron-pair tunnelling, the

effects of which were quantified by Brian Josephson in 1962. He showed that if there is a phase difference between the two wave-functions then a current will flow in the absence of any potential difference.

An insulating layer is not the only possible type of 'weak link', others being a finely ground point contact between the two superconductors or a microbridge formed by etching a superconducting thin film. In practice, a current can be fed at zero voltage through the junction up to a threshold value above which a voltage appears across the junction. The voltage then increases with increasing current. This phenomenon is known as the 'd.c. Josephson effect'. The 'a.c.' effect occurs when a voltage is applied across the junction and oscillatory currents are found to flow.

As will be seen, the Josephson effect permits many applications. But it can also be an irritant. It arises at grain boundaries within polycrystalline samples of the new oxide superconductors and, for example, hinders attempts to measure London penetration depths.

Materials

Of the chemical elements, only non-magnetic metals are found to superconduct, and none of them has a critical temperature greater than 10 K. The addition of impurities generally has little effect on the transition temperatures but, as we have seen, may drastically affect the magnetic and current-carrying properties. Metallic compounds may also be superconducting, and those compounds of vanadium and niobium with the cubic β -W(A15) structure are found to be particularly prone. These had the highest critical temperatures (up to 23 K) before the new oxide variety was discovered. The highest critical fields, on the other hand, were associated with the 'Chevrel phase' compounds with the general formula $\text{M Mo}_6\text{X}_8$. Mention should also be made of 'exotic' species such as organic and heavy-electron superconductors. As with the oxide materials, their mechanisms are not well understood.

Applications

Computers. Josephson junctions offer the capacity to switch ten times faster than semiconducting circuits, with energy losses several orders of magnitude smaller. The latter aspect clearly allows much higher packing density.

Logic circuits exploit the transition between a zero and a finite voltage at a particular current threshold. A Josephson junction is put in parallel with a resistive circuit and couple to three control lines. Depending on the combination of signals in the control lines, a voltage may or may not be generated in the resistance. Thus AND, OR and other logic arithmetic

can be achieved.

Memory circuits exploit the capacity of two junctions in a loop to carry persistent currents in the absence of an applied voltage, in switchable clockwise or anticlockwise directions that can be made to correspond to the values 1 or 0.

Microwave and sub-millimetre detectors.

These use the highly nonlinear current-voltage relationship of Josephson junctions to mix the output of a local oscillator with an incoming signal, as in a standard heterodyne receiver, to produce a signal at an intermediate frequency. If the oscillator current flowing through the junction has a frequency only slightly different from that of the much weaker signal, the resulting current will be amplitude-modulated by an amount proportional to the signal amplitude.

SQUIDS (superconducting quantum interference devices).

These devices provide the most sensitive instruments available for the measurement of magnetic fields, magnetic susceptibilities and voltages. The d.c. SQUID consists of a superconducting ring containing two Josephson junctions. A current which exceeds the zero-voltage threshold of both junctions is pulsed from one side of the ring to the other. Due to quantum interference between macroscopic wave-functions, an imposed magnetic field will change the voltage generated to an extent that is periodic with the field strength, with one cycle per fluxoid, allowing exceedingly sensitive detection (changes of 10^{-11} teslas with a 1-cm ring). A radio frequency (r.f.) SQUID consists of a ring with a single junction coupled to a tuned r.f. circuit. Both d.c. and r.f. SQUIDS have been constructed with high-temperature materials, exploiting their granularity to provide intrinsic Josephson junctions.

Magnets. Superconductors enable magnetic fields to be generated with zero dissipation and, indeed, are now routinely used when very high fields are required. Desirable characteristics of the material are a high transition temperature (which ensures a high upper critical field), high normal-state resistivity (which leads to a large penetration depth compared to the coherence length, and which in turn ensures a higher upper critical field), and high disorder to enhance flux pinning and hence increase the critical currents. Such high-resistance material has the disadvantage of generating heat if it becomes normally conducting, so windings are usually clad in copper to provide a shunt. Unfortunately, the advantages of high critical fields are, for some metal compounds, offset by their brittleness. This also happens to be but one of the several problems that has impeded the use of the high-temperature superconductors. □

Philip Campbell is Physical Sciences Editor of Nature.

Earthquake damage

SIR—Tiedemann¹ has recently commented on our paper² on the resonance effects of the September 1985 earthquakes in Mexico City. Some of his comments and proposed 'rules' are quite relevant and deserve a detailed answer. However, in his opening statement he implies that the distribution of damage simply follows the geographical distribution of 5–15 storey buildings, thus implying that our conclusions are invalid. This is not true. Conservatively speaking, over two-thirds of the 5–15 storey structures were virtually unscathed and even within the lake region of Mexico City there are whole areas with a high density of such structures that were only lightly hit. For example, the Cuauhtemoc area has a larger concentration of ten storey buildings than Roma and there was not nearly as much damage. Furthermore, areas with no high buildings but with "small, old and vulnerable structures" did sometimes suffer great destruction. One case is the Tepito area where such buildings collapsed in great number. (Tepito corresponds to letter B in ref. 2, Fig. 1). We predicted this to be the centre of highest acceleration (see Fig. 2 in ref. 2).

Tiedemann does raise some pertinent points about our paper. The eastern limits of the extinct lake are not well defined, as we emphasized¹, but recently the proposal of a sunken mountain range (Sierrita del Peñón), that essentially coincides with our proposed limits, has been forwarded³. We are aware that the composition of the different deposit layers is not homogeneous and this has to be taken into account; we are currently working on this. The contrast between the soft-ground area and the surrounding volcanoes and hills is, however, far greater than the changes within the former which should only provide refining factors.

It is indeed deplorable that so few accelerographs were available. The only apparatus located within the ancient lake area (letter G of ref. 2, Fig. 1) shows a striking dominance of the 2-s period (ref. 2, Fig. 2).

We also agree with Tiedemann that earthquake damage is controlled by many parameters. As a consequence, all new ideas should be carefully examined and not hurriedly discarded. We insist that we do not imply that we can explain the damage pattern only by the interference pattern. This impression may have stemmed from some comments on our work that overemphasize our statements (see for instance ref. 4). In ref. 2 we report that both pressure and shear waves are transmitted at the old lake's interphase and while the former produce interference patterns the latter cover the whole area uniformly, adding significantly to the tremor effects. In any case our present

model gives a useful first approximation to the interference pattern. Furthermore, we are now investigating a more realistic model² that will predict the directions of ground motion within the relevant area, a prediction that may be tested in future earthquakes for which a much more reasonable amount of seismographic data should be available. At a time when human lives are dependent on construction and protection of buildings, the neglect of any fact (theoretical or empirical) is, as Tiedemann states, dangerous.

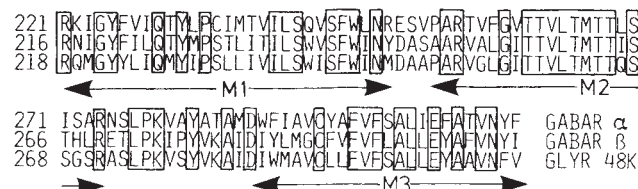
J. FLORES
O. NOVARO
T.H. SELIGMAN

*Instituto de Física,
UNAM, AP20-364,
Mexico 01000 DF, Mexico*

1. Tiedemann, H. *Nature* **329**, 677 (1987).
2. Flores, J., Novaro, O. & Seligman, T.H. *Nature* **326**, 783–785 (1987).
3. Alvarez, R. *Engineering* **1**, (abstr. 1.5) 107–108 (1986).
4. *New Scientist* **28**, 7 May (1987).

Glycine vs GABA receptors

SIR—Comparison of the primary structures of the strychnine-binding subunit of the glycine receptor¹ and both subunits of the GABA_A receptor² has revealed the existence of a gene family for neurotransmitter-gated ion channel receptors which comprises both anionic and cationic chan-



Alignment of transmembrane domains M1 to M3 of the GABA and glycine receptor polypeptides. Identical amino-acid residues are boxed. The sequences include the nicotinic acetylcholine receptor (nAChR) and presumably those gated by amino acids involved in excitatory neurotransmission, such as glutamate. Sequence comparison shows that the subunits of all known members of this gene family have the same number and distribution of membrane-spanning helices, resulting in an identical transmembrane topology^{1–3}. The first transmembrane region predicted, M1, contains a proline residue at an identical position pointing to a common mechanism of channel gating^{1,2}.

In addition, the subunits of the glycine and GABA_A receptor display some striking structural similarities. First, their sequences are considerably more similar to each other (34%–38%) than they are to the subunits of nAChRs (15%–20%). Second, the GABA_A and glycine receptors are most similar within the first three membrane-spanning regions (see figure); eight contiguous amino-acid

residues (TTVLTMTT) are invariant within M2. As segment M2 is thought to line the channel lumen of the different ion-conducting neurotransmitter receptors^{1,2,4–6} we suggest that this conserved sequence will be characteristic of all ligand-gated chloride channels.

Third, a large number of positively-charged amino-acid residues occurs in the putative chloride channel-mouth domains of both glycine and GABA_A receptor polypeptides. These clustered charges have been implicated in anion binding and channel selectivity^{1,2} and may constitute a general feature of ligand-gated anion channels.

Finally, chloride-conducting neurotransmitter receptor proteins have no amphipathic α -helix preceding the fourth hydrophobic region, M4. Such an amphipathic helix is found in all nAChR subunits, and was originally proposed to form the inner wall of the cation channel⁷. The above and further comparisons within this receptor family have been made in more detail in an article appearing elsewhere⁸.

These conserved features indicate a late separation of GABA_A and glycine receptors during ligand-gated ion-channel evolution. In the case of these inhibitory channels, the acquisition of ion selectivity probably preceded that of agonist specificity. However, domain exchanges by exon shuffling may also have contributed to the generation of the present ion channel receptor family, since cation-conducting GABA receptors and anion-selective nAChRs have been described in different systems⁹. Sequence and structural information on these latter receptors should contribute to our understanding of how diversity was generated within this important super-family.

GABRIELE GRENNINGLOH
ECKART GUNDELFINGER
BERTRAM SCHMITT
HEINRICH BETZ

*Molecular Neuropharmacology Lab,
ZMBH, University of Heidelberg,
6900 Heidelberg, FRG*

MARK G. DARLISON
ERIC A. BARNARD
MRC Molecular Neurobiology Unit, MRC
Centre,
Cambridge CB2 2QH, UK

PETER R. SCHOFIELD
PETER H. SEEBURG
*Laboratory of Molecular
Neuroendocrinology
ZMBH, University of Heidelberg,
6900 Heidelberg, FRG*

1. Grenningloh, G. *et al.* *Nature* **328**, 215–220 (1987).
2. Schofield, P. *et al.* *Nature* **328**, 221–226 (1987).

3. Noda, M. *et al.* *Nature* **305**, 818–823 (1983).
4. Giraudat, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 2719–2723 (1986).
5. Oberthür, W. *et al.* *EMBO J.* **5**, 1815–1819 (1986).
6. Imoto, K. *et al.* *Nature* **324**, 670–674 (1986).
7. Finer-Moore, J. & Stroud, R.M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 155–159 (1984).
8. Barnard, E.A., Darlison, M.G. & Seeburg, P.H. *Trends Neurosci.* **10**, 502–509 (1987).
9. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer Sunderland, Massachusetts, 1984).

Aperiodicity of magnetic reversals?

SIR—Recently, *Nature* has published another analysis of data on the Earth's magnetic reversals where a 'statistically significant period' of $\sim$ of 30 Myr is claimed to have been demonstrated¹. Unfortunately, the statistical argument supporting this conclusion is flawed; the test performed could equally well be taken as demonstrating any other period or no period at all. The argument for significance given by Stothers is based on an incorrect use of a correlation coefficient to compare two columns of figures; the relevant portion of that author's Table 2 is reproduced in the first two columns of the table below. These columns give the peaks of spectra for the 'observed' magnetic reversal series and for an 'ideal' series of events equally spaced 30.1 Myr apart.

Observed series	Ideal series	Rescaled (by 0.8) peaks
48.2	48.0	38.4
38.0	40.5	32.4
30.1	30.1	24.08
25.2	24.0	19.2
21.9	20.7	16.56
19.4	19.5	15.6
17.2	17.3	13.84
15.5	15.1	12.08
13.4	13.1	10.48
12.0	12.0	9.6
10.9	11.0	8.8
10.0	10.0	8.0

To quantify the appearance of agreement between columns one and two, Stothers calculated their correlation as $r=0.997$, and this is the basis for the claimed statistical significance of the period of 30.1 Myr. (Actually, he omits the pair (30.1,30.1) from the computation. This affects neither his result nor the point made here.)

The third column given here is formed by multiplying column two by 0.8; it thus represents peaks in the spectrum for an ideal series of events equally spaced by $30.1 \times 0.8 = 24.08$ Myr. Of course, one of the properties of the correlation coefficient is that it is unchanged by linear rescaling, so the correlation of columns one and three is also $r=0.997$, and, if we accept the argument as presented, we are led to conclude that there is also a 'statistically significant period' of 24.08 Myr. In fact, the number 0.8 is arbitrary; can the magnetic reversal series really be in 'statistically significant' agreement with any (and all) possible periods?

The basic problem is that the correla-

tion coefficient is an inappropriate device for comparing two monotone sequences — any two monotone sequences will appear to be highly correlated. The author has tried to allow for this, but the attempt has misfired. To assess significance he has compared his value 0.997 to what is an irrelevant set of reference values: 5,000 sets of eleven numbers were drawn independently and uniformly from the range 10–58 and placed in increasing order; their correlations with the observed series (omitting 30.1) served as the reference set, and in comparison with this set the value 0.997 does seem extreme. But peaks of spectra of random series do not behave like uniformly distributed random variables, whether they are periodic or not. The peaks exhibit patterns of the type the author's Figure 2 displays: clumping at low values and increasing sparsity at large values. If the author had referred his correlation coefficient to a set of values generated in some reasonable way from spectra of random series, the test might have been of some interest (although as a test of general periodicity, not of a 30.1-Myr period, because of the correlation coefficient's insensitivity to rescaling). As it is, the test is utterly uninformative about the stochastic behaviour of magnetic reversals. As Stothers notes, the other test he performed to bolster the claim of statistical significance is questionable on other grounds (it allows for neither the non-stationary nature of the series nor the *a posteriori* selection of 30 Myr as a period of interest). The question of the existence of a period in these data remains open.

STEPHEN M. STIGLER

Department of Statistics,
University of Chicago,
5734 University Avenue,
Chicago, Illinois 60637, USA

1. Stothers, R. B. *Nature* **322**, 444–446 (1986).

STOTHERS REPLIES—Stigler criticizes two aspects of my formal tests for statistical significance of an apparent 30-Myr periodicity in the Earth's magnetic reversals¹. First, he argues that, as the correlation coefficient between any two sequences of numbers is independent of whatever constant factor multiplies each sequence, the sequence of spectral-peak periods derived for a perfectly periodic (ideal) time series can be arbitrarily multiplied by a constant factor and compared with the analogous sequence derived from an observed time series. Thus, any choice of generating period for the ideal time series must give rise to the same value of the correlation coefficient.

Stigler's argument is largely incorrect for the following reason. Consider the harmonic structure, rather than the absolute values, of the spectral-peak periods. (The high peaks are quite obvious in the spectra; the contrast between them and all

lower peaks is sufficiently large so as not to require any special criteria to distinguish them.) To cover the same range of (dimensionless) periods in the observed and ideal spectra, the periods of the peaks must be normalized so that the generating period is formally equal to the assumed basic observational period. But a different choice of generating period necessarily implies a different number of cycles covered in a time series of fixed length. This change affects both the number and the relative period distribution of the spectral peaks. For example, if the generating period is taken to be (dimensionally) 15.5 Myr, which corresponds to one of the major peaks in the observed spectrum and happens to be the basic period advocated by Mazaud *et al.*², the ideal time series covering the same length of record, 165 Myr, must contain 11 dates (as compared to six dates if the generating period is 30.1 Myr). The spectrum for the ideal time series in this case consists of 14 high peaks over the period range 10–58 Myr, instead of the 12 high peaks obtained with a generating period of 30.1 Myr. The spectrum for the observed time series shows 12 high peaks, whose number and relative period distribution (period ratios) are in close accordance with those for a generating period of 30.1 Myr (ref. 1). The dimensional value of the generating period does matter fundamentally, of course, because the physical units in the problem set the dimensions of all periods. Insofar as any other choice of generating period leads to a round number of six dates in the ideal time series, Stigler's criticism is valid, however. This means that the true generating period may lie anywhere in the range 28–33 Myr. But if the question is simply whether 30.1 Myr, in its role as an actual period of one of the high spectral peaks in the observed time series, is the generating period for the other peaks, the answer is yes; the periods of the other peaks will not work as generating periods.

The second objection raised by Stigler is that even for a random time series the distribution of periods of the high spectral peaks is not completely random but is strongly biased toward shorter periods. He is correct here. His suggested remedy, however, is not practicable because not all random time series produce exactly 12 high spectral peaks over the fixed period range 10–58 Myr and because rejection of nonconforming cases would make the test less random. A practical approach is my original test¹, but using 12 ordered but randomly generated frequencies (rather than periods), because the frequency distribution of spectral peaks for both random and ideal (as well as observed) time series turns out, from inspection of many cases, to be roughly uniform (statistically) over the frequency range in question. The occasional generation of two close frequencies in a random set of fre-

quencies is not unrealistic, because my experience with other geological data sets indicates that such close frequencies can occur in analyses of both observed and ideal time series. Proceeding as in my original paper but now using frequency, I still find that $< 0.1\%$ of the random sets of ordered frequencies show a correlation coefficient exceeding the value obtained by using the ordered frequencies derived from an ideal time series having a 30.1-Myr generating period, namely, $r=0.9988$.

Regrettably, Stigler ignores all the other important statistical evidence that favours a basic magnetic period of ~ 30 Myr, such as the occurrence of the highest spectral peak at this period and the basic period's robustness under various processes like truncation, data culling, and prewhitening of the record¹. It is possible, that better tests of significance can be devised and implemented in the future.

RICHARD B. STOTHERS

*National Aeronautics and Space
Administration,
Goddard Space Flight Center,
Institute for Space Studies, 2880
Broadway, New York, New York
10025, USA*

1. Stothers, R. B. *Nature* **322**, 444–446 (1986).

2. Mazaud, A., Laj, C., de Sèze, L. & Verosub, K. L. *Nature* **304**, 328–330 (1983).

P. L. MCFADDEN COMMENTS—Stigler¹ claims that Stothers² use of a correlation coefficient between two monotonically increasing sequences is inappropriate as a formal test for the existence of a 30-Myr period in the magnetic reversal sequence. I agree with Stigler on this point and Stothers³ response (regarding the number of cycles covered in a time series of fixed length) does not overcome Stigler's criticism.

Stothers⁴ accepts Stigler's¹ criticism regarding his original Monte Carlo testing, but rejects Stigler's suggested remedy as impracticable. He notes that not all random time series produce the observed structure (12 high spectral peaks) and suggests that "rejection of nonconforming cases would make the test less random". Here Stothers makes a fatal mistake that has permeated the whole question of periodicities in the magnetic reversal and other records, a point that I return to below. Stothers⁴ then chooses, as the basis for further Monte Carlo sampling, a process that cannot mimic the process he is (by implication) attempting to test as his null hypothesis. Thus his result is of little relevance to the problem at hand.

But the validity of this particular test is a red herring. Stothers² uses a filter that seeks out a particular form of structure. By applying this to the magnetic reversal record he finds a structure with a periodicity ~ 30 Myr. He shows that this structure is robust to truncation, prewhitening and so on. There was really no need for the 'formal' test of a periodicity that Stot-

thers² attempted and that Stigler criticized.

The crucial and fundamental question that has not been addressed is "What is the source of this particular structure: has it been externally imposed upon the reversal sequence or could it have arisen merely by random processes?". To test this properly is not simple and requires the following steps.

(1) Form the null hypothesis that the structure is merely a consequence of the random process. This requires a careful and complete specification of what is meant by 'the random process'. Based on the available evidence I would suggest a gamma process with $k \sim 1.3$ (ref. 4) and with the rate λ reducing linearly from ~ 165 Myr ago to zero at ~ 120 Myr ago, and then increasing linearly from zero ~ 83 Myr ago to ~ 5 at the present time. The precise values will vary with each set of random numbers.

(2) Use a random number generator to create, through the designated random process, a set of 'pseudo-reversal sequences'. (Note that it is incorrect to take the present sequence and reorder the intervals in a random fashion — as has been done in some previous testing — because this does not mimic the process under investigation.) Apply the filter to these sequences to see if it can locate structures similar to the one observed in the actual geomagnetic reversal sequence. By similar I mean that it is not necessary to have a periodicity of 30 Myr, one should merely be identifying whether the random process can produce structures of that form.

(3) If the random process does produce such structures then, on each occasion, there will have to be some specific periodicity that appears as the 'highest spectral peak'. At this stage we must move to the working hypothesis that the actual observed reversal sequence is merely a specific realization of the random process for which the structure happens to have a periodicity of 30 Myr. The important question then becomes "how does the norm (in this case Stothers' residual index) of the observed structure compare with the norms in the pseudo sequences when the pseudo sequence has its highest spectral peak at 30 Myr?".

(4) To answer the question from (3) we must reject all those that do not have 30 Myr as their highest spectral peak and look only at the distribution of those that do. This is then the distribution conditional upon the observed structure and will lead to a properly conditioned test. This is why I stated earlier that Stothers³ makes a fatal mistake (shared by many others) in not wanting to reject the nonconforming cases. To perform the conditional test sensibly we must obtain a large enough sample of pseudo sequences with the highest peak at 30 Myr to ensure a sensible definition of the conditional distribution: only then will we have reasonably reliable

95% or 99% confidence limits. Typically, this will require generation of a vast number of pseudo sequences, most of which will be rejected (even if one is pragmatic and accepts a range of say 29.5–30.5 as being 30), before being able to determine sensible confidence limits. This places severe constraints on the random number generator used, and, in particular, one must be careful that one is not merely cycling through a long sequence of apparently random numbers.

If the observed norm exceeds the critical value of the conditional distribution then either the structure is large enough that it must have been externally imposed, or our model for the random process must be severely wrong (either of which is very interesting). Otherwise the information contained within the reversal sequence does not of itself require any further explanation than the null hypothesis. Whatever filter^{2,5–7} (including any future filters suggested) is used to isolate a structure in the sequence, a properly conditioned test must be performed before it can be claimed that explanation of the data requires an externally imposed structure. As yet no properly conditioned test has been performed and so the question of an externally imposed structure remains open.

P. L. MCFADDEN

*Division of Geophysics,
Bureau of Mineral Resources,
GPO Box 378,
Canberra, ACT 2601, Australia*

1. Stigler, S.M. *Nature* **330**, 26 (1987).

2. Stothers, R.B. *Nature* **322**, 444–446 (1986).

3. Stothers, R.B. *Nature* **330**, 26 (1987).

4. McFadden, P.L. & Merrill, R.T. *J. geophys. Res.* **89**, 3354–3362 (1984).

5. Negi, J.G. & Tiwari, R.K. *Geophys. Res. Lett.* **10**, 713–716 (1983).

6. Mazaud, A. *et al.* *Nature* **304**, 328–330 (1983).

7. Pal, P.C. & Creer, K.M. *Nature* **320**, 148–150 (1986).

Optimization problems

SIR—David Bounds' article "New optimization methods from physics and biology"¹ may leave the reader with the false impression that methods based on simulated annealing, neural networks, and genetic algorithms perform significantly better than traditional optimization methods on NP-complete problems such as the travelling salesman problem. Although the article begins with a discussion of NP-complete problems, the methods described in the paper do not solve any of the problems in this class.

The travelling salesman problem serves as an example. The difficult (NP-complete) version of this problem is to find the optimal route traversing a set of points. A fast solution to this would be of extraordinary interest, since it would lead directly to fast solutions to many other puzzles. Unfortunately, the algorithms described in the article do not solve this problem, but rather the far easier problem of finding a near-optimal route. Conven-

tional algorithms are also able to solve this easier version of the problem², probably much more quickly than the methods described in the paper. The conventional approach also produces results that are probably near to optimal³.

I share Bounds' enthusiasm for these novel methods of optimization, but my own experience in applying them has been that they are more impressive in their range of application than in their speed or quality of results. All three approaches are closely related to simple gradient ascent (hill climbing) in that they depend on local improvements in the cost function to indicate progress toward a global optimum. I find it surprising that so many practical problems yield to such a simple attack. Yet, in applications as diverse as arranging circuits on microchips⁴ and pronouncing written text⁵, these methods produce useful results. Like the elephants in the circus, the remarkable thing is not how well they dance, but that they dance at all.

W. DANIEL HILLIS

Thinking Machines Corporation,
245 First Street, Cambridge,
Massachusetts 02142-1214, USA

1. Bounds, D.G. *Nature* **329**, 215–219 (1987).

2. Garey, M.R. & Johnson, D.S. in *Computers and Intractability* (ed. Klee, V.) 130–131 (Freeman, New York, 1979).

3. Christofides, N. *Tech. Rep.-Grad. School Indus. Admin., Carnegie-Mellon Univ.* 6.1 (1976).

4. Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. *Science* **220**, 671 (1983).

5. Sejnowski, T. & Rosenberg, C. *Complex Systems* **1**, 145–168 (1987).

Why was the fleece golden?

SIR—G.J. Smith's explanation of the Golden Fleece (*Nature* **327**, 561; 1987) has a certain attractive quaintness (in common with other purported scientific explanations applied to legends and folklore) but I would like to point out some contrary evidence.

First, many epicuticular waxes contain oleanoic acid¹. Grapes, for example, contain up to 50 per cent oleanoic acid in their wax, so that over a gram of the acid will be consumed by a person eating a bunch of grapes. Thus, even though triterpenic hydroxyacids seem to be fungistatic², they do not seem to be toxic to human beings³. Some recent reports on intoxication by the consumption of huge amounts of fresh apples^{4,5}, which contain the isomeric ursolic acid, refer to deposition of the hydrocarbons and not to toxicity due to the triterpenic compounds.

PATRICK MOYNA
HORACIO HEINZEN

Faculty of Chemistry,
Universidad de la Republica,
Casilla de Correo 1157,
Montevideo, Uruguay

1. Kolattukudy, P.E. (ed.) *Chemistry and Biochemistry of Natural Waxes* (Elsevier, Amsterdam, 1976).

2. Radler, F. & Horn, D.H.S. *Aust. J. Chem.* **18**, 1059 (1965).

3. Kubo, I. & Matsumoto, A. *Experientia* **40**, 937 (1984).

4. Lange, V.J. & Wildgruber, J. *Fette Seifen Anstrich* **78**, 62 (1976).

5. Duboucher, C. *et al. Chest* **90**, 611 (1986).

Jets in supernova 1987A?

SIR—Speckle observations^{1,2} of SN1987A reveal a secondary light source, 2–3 magnitudes dimmer than the primary and 0.057 ± 0.014 arc s away from it. Its luminosity L_2 is thus $\leq 10^{40}$ erg s⁻¹, and it is $(4.6 \pm 1.1) \times 10^{16}$ cm from the supernova itself. This separation, at $\leq 2.6 \times 10^6$ s, after the explosion implies a relativistic velocity of $v_2 \approx (0.6 \pm 0.15)c/\sin \alpha$, where α is the angle between the velocity and the line of sight. With the main energy reservoir of the collapse depleted by neutrino emission, the kinetic energy available to the ejecta cannot exceed $\sim 5 \times 10^{51}$ erg. The mass of the ejecta, m_2 , must therefore be $< 0.01 M_\odot$. We suggest that the observed secondary object is a relativistic jet. A similar conclusion was reached independently by Rees³, but we differ in suggesting that a rapidly rotating core-collapse has resulted in ejection of jet (or jets) along the rotation axis. T. Nakamura and M. Fukugita (preprint) have suggested rapid core rotation for other reasons. A virtue of this model is that it is based on ingredients which are intrinsic to the supernova and does not require any external object.

Numerical simulations⁴ have demonstrated that a rapidly rotating core collapses first along its rotation axis. Collapse continues along the equatorial directions, then bounces along the rotation axis to form jets. In one example, a $1.4 M_\odot$ stellar core modelled as a $\gamma = 3/4$ polytrope produced jets with mass $0.007 M_\odot$, kinetic energy 5×10^{51} erg, internal energy 3.8×10^{51} erg and average velocity 0.6c. A typical density in the jet, at $r \approx 10^7$ cm from the centre, is 10^{12} g cm⁻³. If the same solid angle is maintained, the density of the jet at the star's envelope, $r \approx 10^{13}$ cm, will be 1 g cm^{-2} , and it will easily pass through the envelope, pushing it aside and accreting mass comparable to or less than its own mass. Rotating collapse is most likely to produce two jets. The luminosity ratio between the two will be $(1 + v_2 \cos \alpha)^2 / (1 - v_2 \cos \alpha)^2 \approx 16$, and only the blue-shifted one will be observed.

Matter accelerated near the core will cool down as it moves outwards, with $T \propto R^{-1}$ (R being the size of the matter blob) if it is radiation dominated, or $T \propto R^{-2}$ if it is a polytrope with $\gamma = 5/3$. (If the ejecta were accelerated near the envelope, at $r \approx 10^{13}$ cm, its thermal energy would be comparable to its kinetic energy, and its temperature would be between 10^6 and 10^7 K. The initial X-ray luminosity would exceed the total initial luminosity of the supernova itself, and this is ruled out by observation.) The luminosity after cooling can be estimated at

$$L \approx (4\pi c/\kappa)(RE_{\text{th}}/m_2) \\ \approx (4\pi c/\kappa)(RE_{\text{th}}^{\text{ini}} f_{\text{cool}}/m_2)$$

where $E_{\text{th}}^{\text{ini}}$ is the initial thermal energy

of the matter, $f_{\text{cool}} = E_{\text{th}}/E_{\text{th}}^{\text{ini}}$ is the cooling factor. When the matter crosses the envelope, $R \approx 10^{11}$ cm and we find that $L_2 \approx f_{\text{cool}} 10^{42}$ erg s⁻¹. The adiabatic cooling from $R \approx 10^7$ cm (at $r \approx 10^7$ cm) to $R \approx 10^{11}$ cm (at $r \approx 10^{11}$) is large enough that by the time the ejecta reaches the envelope its intrinsic luminosity will not pose a problem.

The cooled ejecta must be moderately reheated to emit the observed optical radiation. The needed energy can be supplied by radioactive decay of the jet material (M.J. Rees, personal communication), by interactions with the ambient material³, or by radiation beamed from a pulsar which is formed in the core⁵, and still hidden from us by the envelope. The jet pierces a hole in the envelope through which energy can be supplied. Although the pulsar's luminosity⁵, $L_p \approx 3 \times 10^{42} (B_{12}/10^{12} \text{ G})^2 (P/2\text{ms})^{-4}$ erg s⁻¹, is enough to reheat the ejecta, it is not enough to punch a hole in the envelope by itself. A luminosity of $1.5 \times 10^{44} (\Delta\Omega/0.01)(2 \times 10^6 \text{ s}/T)$ erg s⁻¹ is required over a period of T seconds to create an opening with a solid angle $\Delta\Omega$ in the supernova shell.

One might contemplate replacing the jet by the ejection of a fragment of the core using the sling-shot mechanism of the "Collapse, Pursuit and Plunge" model⁶, but tidal interactions will easily destroy any $0.01 M_\odot$ fragment orbiting near $\sim 0.5 M_\odot$ fragments. Furthermore, the typical rotation velocity of the orbiting fragments is much smaller than the needed 0.6c.

Our model leads to two clear predictions. (1) The secondary light source should be receding from the supernova at a constant rate, and the separation by now (early July) should be four times larger than the separation in late March. (2) The emission from the secondary image should be blue-shifted, and one should look for weak, highly shifted lines in the supernova spectrum due to emission from this source.

We thank Martin Rees for helpful comments on an earlier version of this letter.

TSVI PIRAN

Research Institute for Fundamental
Physics,
Kyoto University, Kyoto 66, Japan
and Racah Institute for Physics,
The Hebrew University,
Jerusalem, Israel 91904.

TAKASHI NAKAMURA

Department of Physics,
Kyoto University,
Kyoto 66, Japan

1. Karovska, M., Nisenson, P., Noyes, R. & Papaliolios, C. *IAU Circ.* No. 4382 (1987).

2. Matcher, S.J., Meikle, W.P.S. & Morgan, B. L. *IAU Circ.* No. 4391 (1987) (erratum *IAU Circ.* No. 4369).

3. Rees, M.J. *Nature* **328**, 207, 1987.

4. Nakamura, T. & Sato, H. *Prog. Theor. Phys.* **66**, 2038–2052, (1981).

5. Ostriker, J.P. *Nature* **327**, 287–288 (1987).

6. Ruffini, J.A. & Wheeler, in *Proc. Conf. Space Phys.*, 45–174 (European Space Res. Orgon, Paris, France, 1971).

Heard above the hubbub

R. W. Guillery

Encyclopedia of Neuroscience, Vols I and II. Edited by George Adelman.
Birkhäuser: 1987. Pp. 1,565. \$125, £110, SwFr. 248.

NEUROSCIENCE has been growing up fast. With the appearance of this encyclopaedia we are bound to ask whether the discipline has reached maturity and to enquire about its personality. The two heavy volumes provide useful clues, serving as would a celebration, where the character and the behaviour of the guests tell much about the person being honoured.

The most important and numerous guests at this party are from the Society for Neuroscience, which has for many years acted as a combination of nanny and fairy godmother, and she has brought many of her best friends. At first, it is difficult to see much order in the clamour of about 700 separate entries, each covering a page or two of double-column print. Everyone seems to be talking at once and many are talking about themselves rather than their subject. A quick survey made up of three separate samples, 15 entries per sample, showed mean autocitation indices of 0.4–0.5 (author's citations/total citations). Some of the guests are talking to themselves, not because they choose to but because in neuroscience today no one has much time for listening.

Brodal (the elder) has come dressed as Polonius. He talks wisely about training — "It is apparently not realized ... that knowledge in neuroanatomy ... requires adequate training and experience" — and about a proper sense of the past — "This disregard of previous knowledge reflects an undeserved underestimation of our predecessors". He might be describing any one of many of the entries, but no one is listening, not even the editor and the 32 eminent members of the advisory board.

The major condescension to a historical past is a long, inaccurate list, entitled "Concise Biographies of Contributors to Progress in Neuroscience, 300 BC to AD 1950" and written from one narrow viewpoint. Many minor figures are included but, among others, Marchi, Retzius, Held, Auerbach, Mauthner, Probst, van Gehuchten and Ganser are omitted. Ramon y Cajal is entered as a "... Spanish master of histology. [He] described the retina (1894); applied the Golgi staining technique for myelin and drew (1909) elegant depictions of neurons". The error concerning the Golgi technique is trivial beside the misunderstanding of Cajal's contribution.

The methods of categorization and the choice of entries seem arbitrary. Some entries may have been chosen to give a

particular author a voice, not to cover a particular subject. Evidence for editorial supervision is sparse. Three entries on neuroanatomy show significant overlap. Neurophysiology is not represented. Now that almost everyone is filling cells and drawing them, is neurophysiology thought to be extinct? And does the past really not matter? *Aplysia* has four entries, primates and flies two each, while reptiles, honeybees, leeches, nematodes, octopus and *Paramecium* have one entry each. Cats, rats, mice and frogs must make do with scattered mentions. One entry on primates says almost nothing about the brain and would have done excellently in an encyclopaedia of comparative anatomy. An entry on the "Triune Brain" must have come from a companion volume on neuromythology.

These are superficial impressions. The two volumes include many stimulating and useful essays and cover much important contemporary material. I have used two methods for providing a more detailed evaluation of coverage. One was to use the books to respond to a student's enquiry about the diencephalic or telencephalic origin of the globus pallidus and pulvinar, which was part of a discussion of cell migration in general. The entry on basal ganglia did not help with the globus pallidus: pulvinar is not in the index. When the discussion broadened we were unable to find either the alar plate or the reeler mutant in the index. Later, I

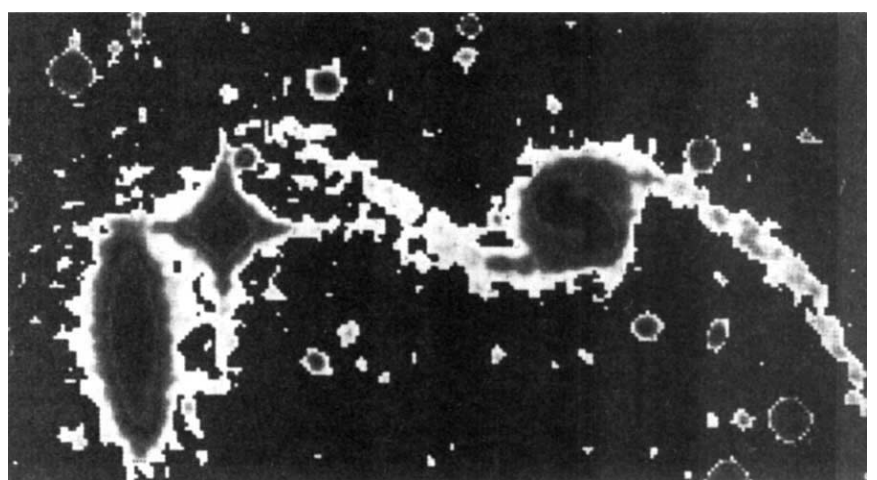
did stumble on this mutant in a fine summary of mutant studies by R. Sidman, and we found that P. Rakic provides an admirably accessible summary of his views on cell migrations and radial glia.

A second approach has been to look at topics I felt I knew well. The entry on the lateral geniculate nucleus was hard to find because the index is poorly organized. When found, it proved to be one of those written for the expert rather than the general reader. The ground rules are assumed, some of the most interesting aspects are omitted and when it comes to detail the entry is wrong in its description of the primate nucleus.

The mamillary bodies make but a brief appearance as a part of the Papez circuit. No caudal connections are mentioned. This may be adequate coverage of an obscure subject. The medulla oblongata is less obscure but just as poorly covered. According to one entry "... in all vertebrates [it] consists primarily of the nuclei of origin and termination of the 5th through 12th cranial nerves". According to another, in reptiles it contains only the 9th to 12th nuclei. Both are wrong. The visual system in contrast, especially the cortex, but also the retina, receives interesting and competent coverage.

One can learn about the neurosurgical treatment of pain, the distributed nature of memory, the receptor concept of drug action, or robotics, and spend hours of fascinating reading. Only the most discriminating readers will be able to distinguish the fine vintage contributions from the ordinary or adulterated ones. Anyone who stays at the party will get a clear sense of excitement about the future, but they'll be lucky if they avoid a hangover. □

R.W. Guillery is Dr Lee's Professor in the Department of Human Anatomy, University of Oxford, Oxford OX1 3QX, UK.



Beyond vision — a computer-enhanced image of two spiral galaxies in gravitational contact. The double system is enormous — IC 5174 (right), which has two long spiral arms, is 750,000 light years from tip to tip, and the galaxies have a radial velocity of 550 million light years. The picture is taken from *Exploring the Southern Sky* by Svend Laustsen, Claus Madsen and Richard M. West, just published by Springer-Verlag, price DM98 (DM128 in 1988).

Communication in silence

Victoria A. Fromkin

What the Hands Reveal About the Brain.

By Howard Poizner, Edward S. Klima and Ursula Bellugi. MIT Press: 1987. Pp. 236. \$25, £22.50.

FOR at least 2,000 years, philosophers and scientists have attempted to understand the nature of the brain, the nature of the mind, and the relationship between the two. A continuing reason for studying human language has been the historical assumption that language is a 'mirror of the mind' or that "Speech is the only window through which the physiologist can view the cerebral life", as was suggested by Fournier in 1887.

The quotation from Fournier refers to speech because there has been a persistent, though incorrect, view which equates speech with language. Speech (production and perception) is behaviour, the use or performance of those who know a spoken language. Language is the abstract mental cognitive system which permits one to speak and understand. Language also underlies the ability of a deaf person to 'sign' and visually to perceive and understand the gestures of a signing person.

To equate speech with language is to obscure the nature of the linguistic systems which form the bases for all spoken languages and all the sign languages used by communities of deaf people throughout the world. As long as researchers concerned themselves only with spoken languages, there was no way to separate what is essential to the linguistic cognitive system from the constraints imposed, productively and perceptually, by the auditory-vocal modality; that is, there was no way of discovering what is the genetically determined linguistic ability of the human brain.

The research on American Sign Language (ASL; the main language of the deaf in the United States), reported in the book under review, reveals the essential similarities between sign languages and

spoken languages. The authors show that ASL and spoken languages are subject to the same kinds of structural constraints, relate forms and meanings by means of the same kinds of recursive rules, and contain equivalent kinds of sublexical units. It therefore should not be surprising that, following brain damage to the left cerebral hemisphere, deaf signers should show deficits in sign language that parallel the language breakdown in similarly damaged hearing patients. The work discussed in this book, however, is the first carefully controlled investigation of deaf brain-damaged patients.

Poizner, Klima and Bellugi present the results of language and non-language (visual spatial) tests, together with cat-scans of the brain lesions of six deaf signers who suffered strokes to either the right or left cerebral hemispheres. Of major importance is the finding that the left-brain-damaged deaf signers who showed aphasia for sign language retained the capacity to process non-language visual spatial relationships, indicating that

the left hemisphere has an innate predisposition for *language* (not speech or the physical ways in which language is expressed). The patients with right-hemisphere lesions, however, showed the reverse pattern, reaffirming that it is the nature of the cognitive system that determines cerebral dominance or lateralization, rather than the perceptual system or the physical attributes of the stimuli being processed.

The book is the result of meticulous research and is a notable contribution to the understanding of brain and language. It provides the first strong evidence that the brain is equipped for language in any modality, and that the kinds of languages which can be acquired are not determined by the motor or perceptual systems but by higher-order brain mechanisms. It will be essential reading for anyone interested in the brain, language and cognition. □

Victoria A. Fromkin is Professor of Linguistics and Dean of the Graduate Division, University of California, Los Angeles, Los Angeles, California 90024, USA.

Genesis of genius

Chen Ning Yang

The Collected Papers of Albert Einstein, Vol. 1, The Early Years: 1879–1902.

Edited by John Stachel. Princeton University Press: 1987. Pp. 433. \$52.50, £32.90. Paperbound English translation by Anna Beck of the documents in Vol. 1 \$22.50, £14.10.

IT HAS been rumoured for quite some time in the physics community that the editors of Einstein's *Collected Papers* have uncovered some hitherto unknown material about Einstein (1879–1955). The rumour turned out to be true and in the first volume of the *Collected Papers*, covering the period 1879–1902, which has just been published by the Princeton University Press, we read some of this fascinating new material, the most important of which is a collection of 51 letters to and from Mileva Maric (1875–1948) who was Einstein's classmate at the ETH (the Swiss Federal Institute of Technology in Zürich) for most of the four years 1896–1900 that Einstein studied there, and who later became Einstein's first wife, 1903–1919.

The volume is handsomely printed and consists of 142 annotated documents, nine editorial notes, a long (19 pages) annotated excerpt from Maja Einstein's 1924 draft of a "biographical sketch" of Einstein (Maja being Einstein's sister), a very useful collection of short biographies about some of the people whose names appear in the documents, and other supplementary material. It is clear that the editors intended this series of volumes,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Scientist and philistine — Albert Einstein with his sister, Maja.

projected to number around 30, to be the definitive source of historical material about Einstein. If this first volume is the norm, they will succeed.

In this first volume, much is revealed about the young Einstein. A few examples will suffice. He was referred to by Maric thus: "my sweetheart has a very wicked tongue" (document 125), a fact borne out by Einstein's various remarks in several letters included in the volume. Einstein had great difficulty finding a job after his *diplom* at the ETH in 1900, which he attributed to poor reference letters from Professor H.F. Weber, his teacher

Autumn books

Nature's Autumn Books supplement appears in two weeks' time, in the issue of 19 November. The books to be reviewed include *The Oxford Companion to the Mind* edited by R.L. Gregory, *Academic Freedom & Apartheid* by Peter Ucko, *An Urchin in the Storm* by S.J. Gould, *The Omega Point* by John Gribbin, *A View of the Sea* by Henry Stommel, *The Purpose of Forests* by Jack Westoby and Edward Harrison's *Darkness at Night*. Among the reviewers are Conor Cruise O'Brien, William Press, Philip Kitcher, Michael Berry, Owen Gingerich and John C. Marshall.

(document 94), and which Maric attributed, partly, to Einstein's "wicked tongue" and his being a Jew (document 125). He had very definite ideas about what he liked about Maric: "When you are my dear little wife, we will zealously do scientific work together, so as not to become old philistines, right? My sister seemed to me so philistine. You must never become like that, it would be awful for me. You must always remain my witch and my street urchin" (document 131). Future volumes presumably will reveal how his ideas changed later.

As is well known, Einstein literally burst upon the world of physics in 1905 when, as a 26-year-old clerk in the Swiss Patent office, he wrote three revolutionary papers in three widely different and important fields of physics: electromagnetism, quantum theory and statistical theory. What had channelled him into these fields? Was he influenced by his teachers? By his friends? These are longstanding puzzles that have never been solved. The present volume adds new information to these puzzles, but if anything deepens the mystery: It is obvious from his letters that Einstein was already interested in the aether and the electrodynamics of moving bodies as early as 1895 (document 5) and 1899 (document 52), in Planck's quantum as early as 1901 (document 96) and in Boltzmann's molecular theory as early as 1900 (document 75). All these are truly astonishing since there was not the slightest indication that this young student of physics had gotten into these deep problems from discussions with more experienced physicists. If later volumes do not provide new insight on this matter we will be forced to conclude that Einstein was a genius with singularly profound perception and that even before the age of 20 he was a fiercely independent thinker capable of focusing on the really fundamental problems of physics.

From the above it is obvious that I regard the present volume as a great contribution to our knowledge about Einstein. I do have, however, a complaint. The volume addresses itself to readers who are thoroughly bilingual in English and German. For a person who reads only English the volume is very frustrating. To be sure, an accompanying volume provides English translations of the documents and of Maja Einstein's draft biography of Albert Einstein. However, untranslated German quotations pop up in the editorial notes, the annotations and the short biographies so often that they become positively irritating. Would it be too much to ask that future volumes add English translations to all quotations in German? □

Chen Ning Yang is Einstein Professor of Physics and Director of the Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11794, USA.

Oral tradition

Chris Quigg

Concepts of Particle Physics, Vols 1 and 2. By Kurt Gottfried and Victor F. Weisskopf. *Oxford University Press: 1984 and 1987. Vol. 1 pp. 204; pbk \$13.95, £9.95; Vol. 2 pp. 432; hbk \$45, £40.*

REVOLUTIONARY progress in particle physics over the past two decades has given birth to a radically new and simple view of the fundamental constituents of matter and the interactions among them. Quarks and leptons, at least 1,000 times smaller than the nucleus, if not truly structureless and indivisible, have been identified as the elementary particles for this generation of scientists, and perhaps beyond. The four fundamental interactions — strong, weak, electromagnetic and gravitational — owe their form to symmetries, expressed in the mathematical language of gauge theories. A common language holds out the prospect of unification of the forces of Nature, extending the achievement of Maxwell and, in the popular mind, surpassing the dream of Einstein. The resulting paradigm embraces a staggering richness of phenomena. It is conceptually simple, but entails a daunting sophistication of theoretical tools, and a remoteness from common experience.

This dramatic progress raises challenges that are shared by other rapidly developing disciplines: how to communicate the essence of the field to non-specialists, how to propagate a teacher's wisdom to beginning students without forcing them to penetrate a thicket of detail. The solution proposed to us in these volumes by two gifted teachers lies in the telling of tales, in inferences by analogy, in intuitive modes of thought communicated by word of mouth from scientific parent to child. Aiming for a wider dissemination of the oral tradition of physics by means of the written word, Gottfried and Weisskopf have set down the substance of lectures given over a period of several years to summer students at CERN, the European laboratory for particle physics.

The result is not a traditional textbook: there are no exercises, not many detailed calculations and very few references from which the reader might learn more. Foregoing a comprehensive bibliography may preserve some of the immediacy of the lectures from which the book developed, but it is very frustrating in a work of this length.

In contrast to the absence of detailed references or suggested readings, there are footnotes on nearly every page containing qualifications and elaborations of the textual material. I found these

distracting; in many cases they contain remarks too detailed to interest the reader for whom the book is apparently intended. Some footnotes are uninformative: the remark that time reversal is an exception to the rule that symmetry operations are represented by unitary operators would be far more useful if it contained the information that time reversal is represented by an *antiunitary* operator, and a definition. Elsewhere, the text would have benefited from rewriting to eliminate ambiguity or to incorporate the subtle thoughts now consigned to footnotes.

The first volume is devoted to a 170-page introduction to the basic concepts and phenomena of particle physics. Brief treatments of prehistory, nonrelativistic quantum mechanics and atomic physics, relativistic quantum mechanics and nuclear phenomena precede the main agenda. All the principal topics are touched upon: the spectroscopy of the strongly interacting particles, their quantum numbers and the related global symmetries; quarks and leptons; and the phenomenology of the strong, weak and electromagnetic interactions. While the selection of topics is apt, much of this presentation is imprecise or muddled, and some of it is simply wrong. It is not correct, as Fig. 12 would have it, that the ϕ -meson is an SU(3) flavour singlet; it is a bound state of strange quark and antiquark, which is as nonsinglet as can be. There are many similar slips of the pen, as well as examples of faulty logic. Although this volume does give an engaging overview of the subject and a hint of current thinking and current issues, it is disappointing that it is not more consistently reliable. I can imagine giving it to a student who has already assimilated Perkins's *Introduction to Particle Physics* for a rainy day read, but not for close study.

The second volume revisits the subjects treated in the first in more detail and at a deeper level. It is generally more successful, but presumably is intended as counterpoint to a standard treatment, rather than as a plausible replacement. An introductory chapter on quantum electrodynamics introduces the authors' notations and summarizes a number of experimental tests of QED, but gives less emphasis to the decisive role of gauge invariance than it merits in the modern view. A descriptive chapter on the static properties of hadrons might well have been incorporated into Vol 1. Quantum chromodynamics, the theory of the strong interactions of quarks and gluons, is developed from the point of view of colour-electric and colour-magnetic field strengths. The treatment of antiscreening by an electromagnetic analogy is not standard textbook fare, so it is pleasing to see it worked out here. The impact is

weakened by relegating part of the punch line to an appendix, however. A discussion of the bag model is murky and out of date, and raises, but does not settle, the question of the η and η' masses. The chapter on deeply inelastic lepton-hadron scattering, the source of much of our experimental information on the structure of the proton, suffers from inconsistent, non-standard notation and ancient data. The essential ideas are here, but in a rather graceless form. Notational to and fro that might seem spontaneous on a blackboard appears simply indecisive in print.

The treatment of electroweak interactions analyses the phenomenology of neutral current interactions and properties of gauge bosons before coming to terms with the spontaneous breaking of the electroweak gauge symmetry. Notwithstanding an insightful discussion of the recognition of the gauge symmetry, this chapter does not entirely succeed. When intuitive arguments and elementary techniques lead without noticeable effort to insights about profound and challenging problems, it is wondrous to behold. But when the approach comes up short, it simply calls attention to the self-imposed handicaps and interferes with the subject itself. Again there is all too frequently an annoying imprecision: a charge

asymmetry in the reaction $e^+e^- \rightarrow \mu^+\mu^-$ is not a parity-violating effect, but only evidence for some mechanism besides single-photon exchange. The oral tradition must stand for impeccable standards of scholarship, as well as insight. When at last it appears, the treatment of hidden symmetries contains some instructive remarks on the analogy with superconductivity and ferromagnetism, and is worth reading.

Overall, there is a disappointing inattention to detail in the production of the book. Assiduous copyediting should have removed awkward turns of phrase, disagreements between subject and verb, and artefacts from the original manuscript. Typographical errors are common enough to be intrusive.

Despite the authors' contagious love of the subject and their desire to communicate the excitement and progress of recent years to a broad audience, *Concepts of Particle Physics* is no royal road to wisdom. Viewed not as a textbook to be studied, but as an anthology of lectures to be browsed, it has a place in departmental libraries. □

Chris Quigg is Deputy Director for Operations of the Superconducting Super Collider Central Design Group, Lawrence Berkeley Laboratory, One Cyclotron Road, Berkeley, California 94720, USA, and Visiting Professor in the Department of Physics, University of California at Berkeley.

No room for loners

Robert M. May & Naomi E. Pierce

Animal Societies: Theories and Facts. Edited by Yosiaki Itô, Jerram L. Brown and Jiro Kikkawa. *Japan Scientific Societies Press*, 2-10 Hongo, 6-chome, Bunkyo-ku, Tokyo 113, Japan: 1987. Pp. 291. ¥8,200.

FROM 1983 to 1986, the Japanese Ministry of Education, Science and Culture sponsored a special research project on "Biological Aspects of Optimal Strategy and Social Structure". This project, directed by Ei Teramoto, built on traditional Japanese strengths in ethology and population biology to fund more than 100 behavioural ecologists. The enterprise was the second largest ever undertaken by Japanese ecologists (the largest being the IBP project in the 1960s), and culminated in an international symposium in Kyoto in July 1986. This book represents the proceedings of the symposium.

Because biologists all over Japan were invited to apply for funding from the project, it is not surprising that a wide range of topics was studied, within the broadly unifying theme of the evolution of social behaviour. The book thus contains a rather eclectic collection of papers, with contributions on wasps, aphids and

dragonflies alongside discussions of pigmy chimpanzees, raccoon dogs and cattle egrets. These diverse studies have been organized under three headings: "Evolutionary Processes", "Competition and Cooperation" and "Economics of Behavior" — which are deftly woven together in the Introduction by J.L. Brown.

Among the senior contributors, Y. Itô (who organized the symposium) uses meticulous empirical observations to support his discussion of the importance of ecological constraints in the evolution of sociality among pleometrotic, tropical wasps. S.F. Sakagami and Y. Maeta present the results of elegant experiments designed to examine factors that would promote social behaviour in carpenter bees that live at the boundary between a solitary and social life. H. Matsuda provides an interesting mathematical metaphor for the evolution of altruism, in which each lattice point in a one- or two-dimensional lattice represents a habitat which accommodates only one haploid individual. He shows that 'attackers' (defined as genotypes with a tendency to attack individuals at neighbouring lattice points, thus increasing the probability of vacant sites for offspring to colonize) have a short-term advantage, but that — under certain conditions, such as offspring not migrating too far — 'altruists' (defined as genotypes with a tendency to help neighbours) can evolve over the long run,

without any explicit kin recognition.

Animal Societies contains several outstanding examples of the work of younger people who were funded by the project. M. Hiraiwa-Hasegawa's analysis of infanticide in primates draws from comparative studies and from her own research on chimpanzees to give an excellent review of a difficult subject. S. Aoki's discussion of the evolution of sterile soldiers in aphids is a highlight of the book: Aoki first discovered the phenomenon of soldier aphids while he was a graduate student in Hokkaido, and he has gone on to describe about 20 species which show this caste differentiation. In this paper, Aoki has synthesized some of these separate observations into a fascinating comparative study that makes plausible arguments about the evolution of sterile soldiers in the cerataphidine horned aphids.

H. Ubukata discusses the behaviour and mating system of the dragonfly *Cordulia aenea amurensis*, and develops a model to explain the searching behaviour of males (who often patrol in straight lines). Ubukata concludes that at high male densities there is an optimal length of territory, with time being wasted in conflict if longer segments are patrolled; at low male densities, the probability of mating success increases as patrol length increases. These theoretical conclusions agree well with observations. Additional interesting papers were presented by other young Japanese scientists at satellite meetings following the main session, and it is perhaps a pity that some of these were not included in the book.

The overseas visitors to the symposium included M.J. West-Eberhart, A.H. Harcourt, M. Taborsky, J. Kikkawa, J.L. Brown and J.R. Krebs, who also contribute stimulating essays on topics central to the theme of the symposium. Kikkawa's paper is notable because it presents new data on social relationships and fitness from his long-term study of silvereye populations on Heron Island.

In our view, the most engagingly original contribution comes from W.D. Hamilton, who turns a tired subject on its head and muses, not about why animals such as hymenopterans and aphids are social, but about why they are not even more cooperative than they already are: "Why is the number of [aphid] morphs so small, and the degree of cooperation effected so slight, compared to conditions found in the assemblage of cells that make up the metazoan body?". Hamilton concludes on a quirky note, with a futuristic view of the final world organism as a "very self-assured, competent, and doubtless, for this case, female, medical student". For an explanation of this vision, you must see the book. □

Robert M. May and Naomi E. Pierce are in the Biology Department, Princeton University, Princeton, New Jersey 08544, USA.

The brain as a Darwin Machine

William H. Calvin

For parallel computers to simulate our brains, we must face the fact that human beings have a better claim on the title Homo seriaticum than Homo sapiens — we're more consistently serial than wise.

AMIDST all the hyperbole about thinking machines that has accompanied the emergence of large-scale parallel computers from their serial predecessors, we have begun to contemplate the prospect of simulating some of our brain's massive parallelism. But one immediately runs into a role reversal worthy of a Mozart opera: the most distinctively human higher brain functions are surprisingly serial.

Human beings are perpetually stringing things together: phonemes into words, words into sentences, concepts into scenarios — and then fussing about getting them in the right order. Our brain uses word-order rules to create a very productive language, with an infinite number of novel messages, rather than the several dozen standard interpretations associated with the several dozen cries and grunts of any other primate species. It is not our mellifluous voices that constitute a significant advance but rather our arrangement rules, the meaningful order in which we chain our utterances.

Further, talking-to-ourselves consciousness is, among other things, particularly concerned with trying to chain together memory schemata to explain the past and forecast the future. As literary critic Peter Brooks has said¹:

Our lives are ceaselessly intertwined with narrative, with the stories we tell and hear told, those we dream or imagine or would like to tell, all of which are reworked in that story of our own lives that we narrate to ourselves in an episodic, sometimes semiconscious, but virtually uninterrupted monologue. We live immersed in narrative, recounting and re-assessing the meaning of our past actions, anticipating the outcome of our future projects, situating ourselves at the intersection of several stories not yet completed.

It is our ability to choose between such scenarios that constitutes our free will — though, of course, our choices are only as good as our imagination in constructing a wide range of candidate scenarios. Logical reasoning also seems dependent upon the rules of reliable sequencing. Our sophisticated projection abilities are very sequential: a chess master, for example, tends to see each board configuration not just after the next move but a half-dozen moves ahead, as several different scenarios.

Not that you could get much human-

like language or scenario-spinning consciousness out of the ordinary serial computer — what we are probably talking about is parallel architecture being used to create a lot of serial paths from which to choose. And, perverse though it may seem, we are also likely to make inten-

"Technology treats noise as an unwanted impediment, darwinism as a means of exploring new avenues. But here we see it as a stimulus to evolve redundant machinery — whose secondary uses may be revolutionary."

tional use of noise, good old randomness ('stochastic process' is the polite euphemism). I can hear an incredulous voice already: "Not only does he want us to waste our precious parallel computing power simulating an old-fashioned serial device, but he wants us to make our machine intentionally noisy?"

Yet abandoning low-level reliability (and achieving overall reliability via stable superstructures) is very useful. Noise can be creative. Every time that you think of sex, you should remember that it is all about guaranteeing some randomness — shuffling the DNA deck during crossing-over when making sperm and ova. The invention of eukaryotic sex a thousand million years ago probably prompted the great Precambrian diversification of complex life forms into the familiar tree of species.

The brain's construction of chained memories and actions is probably another tree, though a more functional metaphor might be the candelabra-shaped railroad marshalling yard, with words for cars: imagine that many trains are randomly constructed on the parallel tracks, but only the best is selected to be let loose on the 'main track' of consciousness and speech. Best is determined by memories of the fate of somewhat similar sequences in the past, and one presumes a series of selection steps² that shape up candidates into increasingly more realistic sequences. This selection among stochastic sequences is more analogous to the ways of darwinian evolutionary biology than to the 'von Neumann machine' serial computer. One might call it a Darwin Machine³ instead: it shapes up thoughts in milliseconds rather than millennia, and uses innocuous

remembered environments rather than the noxious real-life ones.

Before pursuing such intracerebral Darwin Machines, consider some non-biological examples. Daniel Hillis has been using massive parallelism to create some competing computer programs. They mutate, surviving on the basis of how fast they can put a list of names into alphabetical order. Just using random variations on a basic program loop, his parallel computer has re-discovered many of the known sorting algorithms⁴. Similarly, the artist Harold Cohen's computerized drawing machine AARON makes aesthetically pleasing paintings using random variations and some general selection rules⁵.

Toolmaking can operate the same way, and perhaps did so even two million years ago when hominids had ape-sized brains. The late Glynn Isaac used to demonstrate early toolmaking techniques during his archaeology lectures by pounding together two potato-sized rocks, not delicately but furiously: chips would soon be scattered all over the floor. After a minute, he would stop and sort through the dozens of stone flakes. And he would pick up some excellent analogues of the single-edged razor blade, just the thing for incising the tough hide of a savannah animal, or amputating a leg at the joint. This stochastic toolmaking is one round of a Darwin Machine: make lots of random variants by brute bashing about, then select the good ones. Perhaps another round of bashing resulted in a flake splitting, two sharp edges intersecting in a point. Careful craftsmanship probably developed where the raw materials were scarce.

Darwinian schemes⁶ (some of which have the successive rounds of randomness plus shaping-up selection that mark them as members of the class I am calling Darwin Machines) have also emerged as partial explanations for

- bacterial food-finding⁷ (*E. coli* intermittently randomize their swimming path by tumbling, but suppress tumbling when the nutrient concentration is increasing, and so select random paths that lead towards high concentrations);

- long-term memory consolidation⁸ (selective retention, during a culling procedure, of randomly generated but subsequently used synapses, analogous to

photographic development retaining exposed silver grains⁸); and

- perceptual categorization via shaping up cortical interconnections (Gerald Edelman⁹) notes that, in consequence of the random element, "we must look at all acts of perception as acts of creativity"). Given that half the cortical synapses are disconnected during childhood¹⁰, there is again much opportunity for darwinian editing.

And, for at least a century¹¹, it has been recognized that even the highest-known biological function, human thought, involves random generation of many alternatives and is only shaped up into something of quality by a series of selections. Like the elegant eyes and ears produced by biological randomness, the Darwin Machine's final product (whether sentence or scenario, algorithm or allegory) no longer appears random because of many millisecond-long generations of selection shaping up alternative sequences off-line.

The Darwin Machines of particular interest here are the ones associated with chaining together actions (sequencing). Although they are often useful, command queues for detailed preplanning are seldom essential: goal-plus-feedback usually suffices, as when raising a cup to one's lips and getting progress reports from the joints and muscles. Where planned chains become essential, and thus likely to evolve rapidly, is where feedback becomes impossible, yet a linked series of moves must be precisely executed. Reaction time becomes a problem in brief ballistic movements such as hammering, throwing, clubbing or kicking: the progress reports will usually arrive too late for corrections to be made. For organisms that need to be both large (metres of conduction distance) and fast, one often needs the neural equivalent of an old-fashioned roll for a player piano. During 'get set', we carefully plan to act without feedback.

Grammar

Sequencing may involve much of the left cerebral hemisphere in mammals¹². Left premotor cortex tends to program linked movements for not only the right hand and arm, but the left as well¹³. Left hemisphere is best at deciphering rapid sound sequences — and so it apparently became a natural home for many language-related abilities. Indeed, the core of human language cortex is a sequencing area for both incoming sounds and outgoing movements¹⁴, just what grammar needs.

Comparison of grammars shows that the typical subject-verb-object word order of an English sentence is not biologically determined: Japanese syntax uses subject-object-verb, while classical Arabic puts the verb first. What the biol-

ogy may provide is the serial buffer to hold the phrase while it is analysed according to the learned rules (though more subtle grammatical linkages are perhaps constrained by buffer branchings, corresponding to Chomsky's deep structure). There are some suggestions that the capacity of one important serial buffer is about a half-dozen items, judging from phenomena such as chunking of memory¹⁵.

Yet there is surely more than one serial buffer: the human brain seems to orchestrate many sequences in parallel. Most are subconscious, with only one entering our 'main line' awareness (as in the railroad yard's parallel-to-serial bottleneck); traditional lines of evidence are scene-shifting in dreams, subconscious problem-solving and how subconscious scenarios sometimes pathologically intrude into speech.

Another suggestion of parallel sequencers has arisen from a very different direction. Human beings often hunt with projectiles; faster and farther throws are always better, provided accuracy can be maintained. A biophysical model for throwing¹⁶ has emphasized the need for sub-millisecond timing precision, far in excess of what one can expect from noisy neurons¹⁷ in a single command buffer. This suggests that the precision of the release of a projectile must arise from the Law of Large Numbers (the same rationale as why, in order to halve a standard deviation, one averages four times as much data).

Thus there must be many sequencers which, at least temporarily, can be ganged in parallel (imagine multiple columns of horses pulling a single wagon) during the occasions demanding peak performance in one-shot timing. To hit a rabbit-sized target reliably from twice the distance requires that the jitter in rock release time must be narrowed by a factor of eight, and the only known way of accomplishing this feat is, as one gets set to throw each time, to assign 64 times as many noisy neurons to the task and then average their recommendations for the release time.

Technology treats noise as an unwanted impediment, darwinism as a means of exploring new avenues. But here we see it as a stimulus to evolve redundant machinery — whose secondary uses may be revolutionary. There may even have been a 'noise window' in hominid evolution: lacking sufficient neuron noise to overcome, Ice Age hominids might have become proficient projectile predators without the massively serial scheme¹⁸. While timing precision is the argument for why so many parallel planning tracks were evolved in the first place, the really interesting things are the possible spare-time uses — if those extra buffers are capable of randomly sequencing other things when not needed for throwing-hammering-clubbing muscle commands.

If the separate tracks can also be unhitched to operate independently, then one might expect a Darwin Machine to emerge. By providing many candidate queues, it might foster stringing words together into more sophisticated sentences, or schemata into more credible scenarios. Rather than our productive language and planning-for-the-future consciousness arising gradually through their own selective advantages, they could have emerged as novel spare-time uses of neural machinery originally under selection for more mundane forelimb movements — much as a novelty called bird flight probably emerged willy-nilly as a consequence of natural selection for keeping warm via forelimb feathers (because it takes a lot of feathers to begin flying).

Serendipity

Neural-like networks¹⁹⁻²¹, once they become capable of stochastic generation of sequences, then successive selections by remembered environments, do offer an obvious route to machine intelligence — though, should we succeed, we shall surely have to cope with machine imagination and machine 'free will'. We do not yet know how much of our own mental life might be explained by serendipitous secondary benefits of stochastic sequencers. But just as darwinian gradualism has been supplemented with notions of sexual and group selection, isolation and speciation, stasis and 'fast tracks', so we might expect a fuller understanding of our mental life to identify additional processes that regulate and elaborate the stochastic shaping-up of novel constructs in our intracerebral Darwin Machines. □

1. Brooks, P. *Reading for the Plot 3* (Knopf, New York, 1984).
2. Dawkins, R. *The Blind Watchmaker* (Longman, London, 1986).
3. Calvin, W.H. *Whole Earth Review* **55**, 22-28 (1987).
4. Hillis, W.D. personal communication (1987).
5. McCorduck, P. *Whole Earth Review* **55**, 45-51 (1987).
6. Young, J.Z. *J. R. Soc. Med.* **72**, 801-814 (1979).
7. Segall, J.E., Block, S.M. & Berg, H.C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8,987-8,991 (1986).
8. Calvin, W.H. & Ojemann, G.A. *Inside the Brain* 65-67 (NAL, New York, 1980).
9. Edelman, G.M. in *How We Know* 24 (Nobel Conference, 1985).
10. Rakic, P., Bourgeois, J.-P., Eckenhoof, M.F., Zecevic, N. & Goldman-Rakic, P.S. *Science* **232**, 232-235 (1986).
11. Campbell, D.T. in *The Philosophy of Karl Popper* (ed. Schilpp, P.A.) 413-463 (Open Court, La Salle, Illinois, 1974).
12. Bradshaw, J.L. & Nettleton, N.C. *Behav. Brain Sci.* **4**, 51-91 (1981).
13. Kimura, D. *Phil. Trans. R. Soc. B* **292**, 135-149 (1982).
14. Ojemann, G.A. *Behav. Brain Sci.* **6**, 189-230 (1983).
15. Simon, H.A. *Models of Thought* **41** (Yale University Press, New Haven, 1979).
16. Calvin, W.H. *J. Theoret. Biol.* **104**, 121-135 (1983).
17. Calvin, W.H. & Stevens, C.F. *J. Neurophysiol.* **31**, 574-587 (1968).
18. Calvin, W.H. *The River That Flows Uphill: A Journey from the Big Bang to the Big Brain* 407 (Macmillan, New York, 1986).
19. Rumelhart, D.E., McClelland, J.L. & the PDP Research Group *Parallel Distributed Processing* (MIT Press, 1986).
20. Dehaene, S., Changeux, J.-P. & Nadal, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2,727-2,731 (1987).
21. Kleinfeld, D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9,469-9,473 (1986).

William H. Calvin is in the Biology Program, University of Washington, NJ-15, Seattle, Washington 98195, USA.

North Atlantic thermohaline circulation during the past 20,000 years linked to high-latitude surface temperature

Edward A. Boyle* & Lloyd Keigwin†

* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

During a surface cooling event 10,000 to 12,000 years ago, higher Cd/Ca and lower $^{13}\text{C}/^{12}\text{C}$ ratios are observed in benthic foraminifera shells from rapidly accumulating western North Atlantic sediments. Data from sediment cores show that marked nutrient depletion of intermediate waters occurs in association with reduced glacial North Atlantic Deep Water flux. It is proposed that cold high-latitude sea surface temperatures enhance intermediate-water formation at the expense of deep-water formation.

THE transition from latest Pleistocene glaciation to the Holocene interglacial period occurred over thousands of years¹⁻⁴. In the North Atlantic Ocean and continental Europe, this long-term warming trend was interrupted by a brief return to glacial temperatures ~10,000–12,000 yr ago referred to as the 'Younger Dryas'^{5,6}. The relative abundance of cold- and warm-water fossil fauna in oceanic sediments shows that the surface North Atlantic was also cooler then. Recent accelerator radiocarbon dating of single-species foraminifera from a high-accumulation-rate sediment core indicates a terrestrial-equivalent radiocarbon age of ~10,300 yr⁷. Oxygen isotope studies from ice cores demonstrate that deglacial warming in Greenland was also interrupted by a brief period of near-glacial temperatures⁸. Recent counting of annual $\delta^{18}\text{O}$ and dust cycles indicates that this event ended 8,770 chronological years ago⁹ (equivalent to 9,050 conventional radiocarbon years). These events are synchronous within the uncertainties in the ages; a climate model study indicates that synchronicity is to be expected¹⁰. The sudden reversal of the warming trend is not seen in ice-volume indicators, except perhaps as a pause in melting².

This event is one of the shortest well-documented extreme climate reversals, occurring over $\leq 1,000$ –2,000 yr. The rapidity of this fluctuation can help determine the response time of other climate subsystems. For example, the Younger Dryas can be used to test the response of deep-ocean circulation to a short climate event. The duration of the Younger Dryas climate reversal is short compared to Milankovitch (orbital) cycles, implying that other mechanisms can generate large climatic changes. A number of models have been proposed to account for the Younger Dryas, but no consensus has been reached yet⁶.

No convincing evidence for a deep-ocean response to Younger Dryas cooling has been obtained from high-sedimentation-rate cores. Given normal deposition at 2–3 cm kyr⁻¹ (and typical biological stirring of the upper sediment layer down to 8 cm), it is unlikely that such a short-duration event could be reliably recorded in most deep-sea sediment cores. The north-east Bermuda Rise, a feature created from transport of fine-grained detritus, has high deposition rates, of >10 cm per thousand yr^{11,12}. We present data from a Bermuda Rise core which document the deep-ocean response to Younger Dryas cooling.

The sediment core and its timescale

Core EN120 GGC1 is a large-diameter (10 cm inner diameter) gravity core. It was raised from 4,450 m depth (33°40' N, 57°37' W) on RV *Endeavor* cruise 120, 31 August 1984. The core site is near that of core KNR31 GPC5 for which previous stratigraphic work has been undertaken^{11,12}.

The timescale for this core is estimated by cross-correlation

with features that have been radiocarbon dated in other cores from the North Atlantic basin. This correlation assumes that ice volume is the dominant factor controlling changes in benthic foraminiferal $\delta^{18}\text{O}$, with a lesser contribution from deep-ocean temperature¹³. The ice-volume signal must be globally synchronous within the mixing time of the ocean (<1,000 yr¹⁴); part of the temperature contribution to $\delta^{18}\text{O}$ variability is likely to be correlated with ice volume as well (W. F. Ruddiman, in preparation). The estimated timescale is based on the following assumptions. (1) The sample with most positive $\delta^{18}\text{O}$ is assigned the conventional glacial maximum age of 18,000 yr. Because benthic foraminifera are scarce in the glacial section of the core, there is some uncertainty as to the true depth of the glacial maximum (Fig. 1). The following discussion will not rely on the details of the chronology preceding deglaciation. (2) The first rapid shift of ^{18}O towards interglacial values ('Termination Ia') is very clear in this core at 145 cm, with an uncertainty of no more than a few cm. This feature has been dated at 14,000 radiocarbon years using the accelerator radiocarbon technique in another North Atlantic sediment core¹⁵. (3) After a brief $\delta^{18}\text{O}$ plateau, a second rapid shift to full interglacial conditions (Termination Ib) occurs which has been accelerator radiocarbon dated at 8,500 radiocarbon years¹⁵. This transition is less clearly defined than Termination Ia, but it occurs somewhere between 80 and 95 cm in EN120 GGC1. The Termination Ib level was chosen at 90 cm. (4) There is a regional CaCO_3 minimum (at 17 cm depth in this core), as well as several other CaCO_3 cycles in the Holocene section of this core (Fig. 1), that are identical to cycles observed in nearby core GPC5 (L. D. Keigwin and G. A. Jones, in preparation). Planktonic foraminiferal carbonate from the youngest carbonate minimum has been radiocarbon dated at 1,900 yr in KNR31 GPC5¹¹. (5) The core top is assigned an age of 0 years. In view of the characteristics of the core, similarity of the chemical data for the core top and Smith-McIntyre box-core sample, and the high sedimentation rate, this assignment appears reasonable and is unlikely to be low by more than a few hundred years.

With these control points, other depth intervals are assigned ages by linear interpolation (and extrapolation beyond the glacial maximum). Preliminary results from accelerator radiocarbon dates in this core and nearby core GPC5 support the timescale used here. In particular, in GPC5, the Younger Dryas interval is bracketed by radiocarbon dates of $9,210 \pm 150$ yr (*G. ruber*) at 175 cm and $11,150 \pm 310$ years (*N. pachyderma*) at 202 cm (L. D. Keigwin and G. Jones, in preparation). The CaCO_3 cross-correlation indicates that these intervals in GPC5 are equivalent to 100 cm (interpolated age 9,500 yr) and 115 cm (interpolated age 11,000 yr) in EN120 GGC1. It is unlikely that

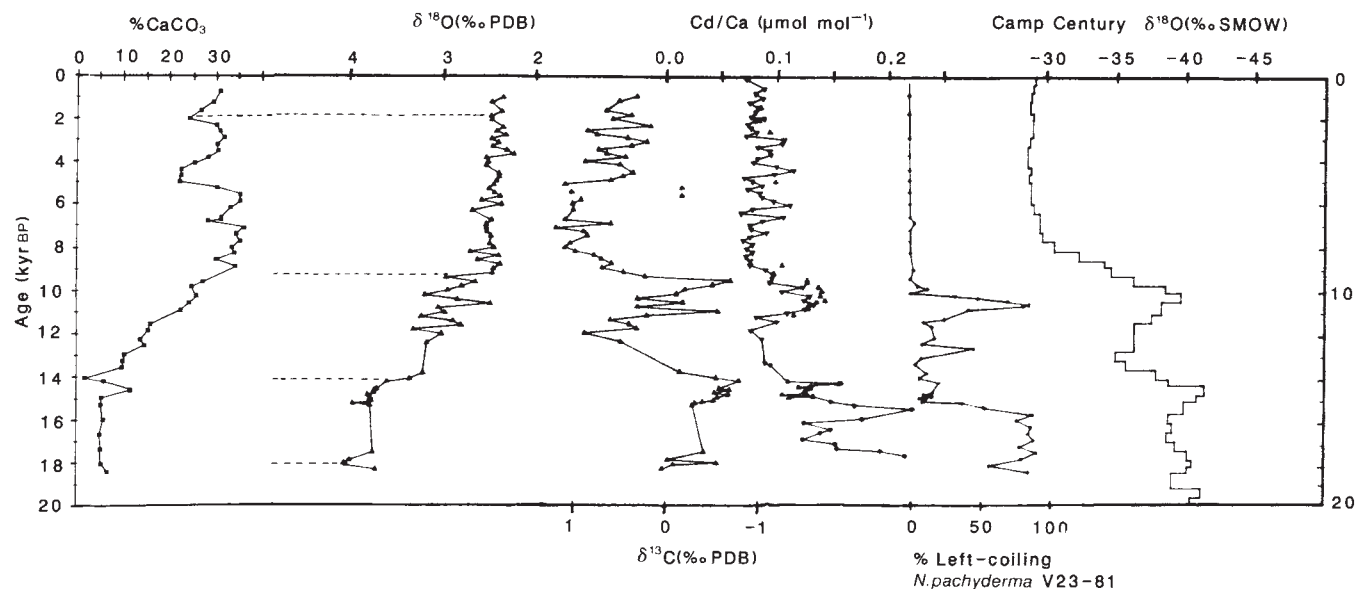


Fig. 1 Data from core EN1200 GGC1 against estimated age. The line connecting $\delta^{13}\text{C}$ points is interrupted where two outlier points occur. In the Holocene section of the core, *N. umbonifera* (inverted triangles) was used for Cd/Ca analyses; in the Younger Dryas section both *C. wuellerstorfi* (triangles) and *N. umbonifera* were used; during the deglacial transition, only *C. wuellerstorfi* was used, and in the glacial period *Uvigerina* (diamonds) was used. Although the use of multiple species might raise concerns about possible differential species fractionation, the analysis of several hundred paired samples of these species shows no evidence for differential fractionation, including intervals of this core where species overlap. In particular, both *C. wuellerstorfi* and *N. umbonifera* were analysed in the Younger Dryas section, and both species confirm the event. The line is drawn through the *N. umbonifera* data, with the *C. wuellerstorfi* data plotted as individual points. Note that the two right-hand datasets are cross-correlated from other records, not measured in our core. $\delta^{13}\text{C}$ ($=1,000 \times ((^{13}\text{C}/^{12}\text{C})_{\text{sample}} - (^{13}\text{C}/^{12}\text{C})_{\text{standard}}) / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}$) is measured in parts per thousand relative to the PDB standard; $\delta^{18}\text{O}$ is measured relative to standard mean ocean water (SMOW).

any of the Holocene age assignments are >500 yr in error. Each 2-cm sample represents the equivalent of 200 years of deposition. If the biologically stirred mixed layer on the sea floor is 8 cm deep (and a gravity-core foreshortening factor of 0.57 is assumed), then the resolution of the record will be better than 400 yr.

Deep-water variability

Benthic oxygen isotope data document the core's continuity of sedimentation and high resolution (Fig. 1). Benthic carbon isotope and cadmium data from this North Atlantic sediment core provide information on deep-water nutrient variability, which is linked to changes in relative proportions of North Atlantic and Antarctic deep-water sources¹⁶⁻¹⁸. Nutrient differences between North Atlantic and Antarctic waters result from the interaction between biological nutrient cycling and the general circulation of the ocean¹⁴. As various sources of deep water spread throughout the ocean, their characteristic chemical signatures remain evident on a global scale¹⁹⁻²¹. Because North Atlantic Deep Water derives from nutrient-depleted waters of subtropical origin (whereas Antarctic waters originate from nutrient-rich waters upwelled from great depth), North Atlantic Deep Water is enriched in ^{13}C and depleted in Cd and nutrients relative to Antarctic waters.

In addition to the effect of chemical redistribution throughout the ocean, chemical inventories change in response to imbalances between input (continental weathering) and output (sedimentation). Changes in the global oceanic inventory will influence the record at each site in the ocean²². Data from ocean sediment cores suggest that glacial mean $\delta^{13}\text{C}$ was $\sim -0.5\%$ relative to the modern ocean, and the oceanic Cd inventory does not appear to change significantly^{22,23}. Carbon isotope variability in this core encompasses equal contributions from changes in the oceanic mean and circulation, while Cd responds mainly to changes in ocean circulation.

As expected, benthic data from this core show that glacial deep water (4,450 m) is depleted in ^{13}C and enriched in Cd relative to the present. This reflects a lower proportion of north-

ern source water relative to admixed southern source water, as seen in previous studies on cores from 2,500–3,500 m^{16,17,22,24}. The amplitude of the signal in this record is larger than that observed for shallower, lower-sedimentation-rate cores. Both $\delta^{13}\text{C}$ and Cd data imply that the most nutrient-depleted waters (hence greatest relative North Atlantic Deep Water flow) occurred in the early Holocene (7,500–9,000 yr ago). There is some hint of upper Holocene fluctuations but these are not equally supported by both $\delta^{13}\text{C}$ and Cd, and hence are not considered here.

Both $\delta^{13}\text{C}$ and Cd data show that an episode of nutrient-enriched deep water occurred between 9,000 and 11,500 years ago. Coincident with (or perhaps preceding) the initial phase of deglaciation 14,000 years ago (Termination Ia), North Atlantic Deep Water proportions rose from a glacial nadir to levels comparable to those found in the modern ocean. These high fluxes continued for $\sim 2,000$ years. Between 9,000 and 11,500 years ago, North Atlantic Deep Water ebbed to a level only slightly higher than that of the glacial maximum. During this event, reduced North Atlantic Deep Water persisted until just before the renewed warming phase (Termination Ib). This ebb of the North Atlantic Deep Water coincided (within the likely error of the timescale) with the Younger Dryas cooling event observed on the European continent, in Greenland, and in high-latitude surface waters of the North Atlantic Ocean. These data demonstrate, for the first time, that this brief cold event is associated with a response of similar import and duration in the deep ocean. It is now evident that the deep ocean can undergo dramatic changes in its circulation regime on timescales of ≤ 500 yr.

North Atlantic Intermediate Water

Higher nutrient concentrations in the deep North Atlantic (2,500–4,500 m) are accompanied by lower nutrient levels at intermediate depths. Down-core analyses of $\delta^{18}\text{O}$, $\delta^{13}\text{C}$ and Cd in benthic foraminifera from western North Atlantic core CHN82 Sta41 Core15PC (43°22.3'N, 28°13.9'W, 2,151 m; Table 2, Fig. 2) do not indicate any increase in nutrient content during

Table 1 Data from sediment core EN120 GGC1

Depth (cm)	Cd species	Cd/Ca ($\mu\text{mol mol}^{-1}$)	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$ (‰ PDB)	Depth (cm)	Cd species	Cd/Ca ($\mu\text{mol mol}^{-1}$)	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$ (‰ PDB)
0	umb	0.072			103	umb	0.123		
	wue	0.073				wue	0.137	2.80	-0.49
5	umb	0.088			105	umb	0.104		
7	umb	0.080				wue	0.140	2.95	-0.19
9	umb	0.087	2.36	0.33	107	umb	0.129		
11	umb	0.075	2.48	0.52		wue	0.139	3.19	-0.10
13	umb	0.082			109	umb	0.124		
	wue	0.085				wue	0.143	2.84	0.32
15	umb	0.080	2.37	0.66	110	wue	0.135		
17	umb	0.076	2.49	0.39	111	umb	0.133		
18	umb	0.088				wue	0.129	2.49	-0.17
19	umb	0.080			113	umb	0.125		
	wue	0.083	2.48	0.59		wue	0.128	3.05	0.32
21	umb	0.073			115	umb	0.109	2.98	-0.55
23	umb	0.076	2.24, 2.48	0.32, 0.03	116	wue	0.115		
25	umb	0.081	2.43	0.86	117	umb	0.081	3.23	0.22
	wue	0.093			119	umb	0.100	2.89	0.61
27	umb	0.072			121			2.80	0.41
	wue	0.127	2.33	0.76	123	umb	0.076	3.31	0.33
29	umb	0.106	2.57, 2.39	0.49, 0.39	125			3.01	0.89
31	umb	0.104	2.41, 2.40	0.53, -0.08	127	umb	0.086		
33	umb	0.082	2.47	0.39		wue	0.111		
35	umb	0.093	2.32	0.74	129			3.17	0.50
37	umb	0.094	2.24	0.65	138	umb	0.089		
39	umb	0.082	2.54	0.45	139	umb	0.094		
41	umb	0.078	2.52	0.88	143			3.21	-0.14
43	umb	0.099	2.54	0.51	147			3.35	-0.54
45	umb	0.114			149	wue	0.110		
47	umb	0.097	2.40	0.37	151	wue	0.158	3.59	-0.78
49	umb	0.070	2.40, 2.40	0.77, 0.19	152	wue	0.156		
51	umb	0.077			153	wue	0.135		
	wue	0.098	2.43	0.61	155	wue	0.131		
53	umb	0.086	2.46	1.10	157	wue	0.120		
55	umb	0.073	2.52, 2.49	-0.34, 0.01	159	wue	0.130		
57	umb	0.084	2.45	1.03	161			3.70	-0.57
59	umb	0.086	2.43, 2.35	-0.18, -0.14	163	wue	0.125	3.73	-0.68
61	umb	0.096	2.59	0.93	165	wue	0.127		
63	umb	0.111	2.37	1.02	167	wue	0.105	3.74	-0.52
65	umb	0.077			169	wue	0.133	3.80	-0.67
66			2.69	1.01	171	wue	0.111		
67	umb	0.086			173			3.77	-0.55
69	umb	0.105			176	Uvi	0.149		
71	umb	0.086	2.48	1.10	177			3.76	-0.51
73	umb	0.075	2.54	0.61	179			3.79	-0.39
75	umb	0.076	2.54	1.20	181	Uvi	0.170	3.96	-0.31
77	umb	0.090	2.53	0.90	183			3.77	-0.28
79	umb	0.075			186	Uvi	0.222		
	wue	0.106	2.49	0.86	199	Uvi	0.177		
81	umb	0.069			204	Uvi	0.125		
83	umb	0.077	2.50	1.05	213	Uvi	0.149		
85	umb	0.072	2.45	1.10	217	Uvi	0.140		
87	umb	0.077	2.71	0.99	226	Uvi	0.124		
89	umb	0.071	2.40	0.79	232	Uvi	0.153		
91	umb	0.075	2.56, 2.69	0.38, 1.04	238	Uvi	0.155		
93	umb	0.075	2.39	0.60	242	Uvi	0.194		
95	umb	0.089	2.46	0.70	245			3.74	-0.41
97	umb	0.095			248	Uvi	0.216		
	wue	0.097	2.47	0.47	256			3.99	-0.02
99	umb	0.095			260			4.05	-0.55
	wue	0.127	2.97	0.24	264			4.02	-0.08
101	umb	0.093			273			3.71	0.04
	wue	0.127	2.65	-0.68					

All isotope data is from *C. wuellerstorfi*. '0'-cm sample is from Smith-McIntyre surface sediment sampler; all other depths refer to distance along core liner (sediment surface in core is at 4 cm). In the species column the symbols are: wue, *C. wuellerstorfi*; umb, *N. umbonifera*; Uvi, *Uvigerina*.

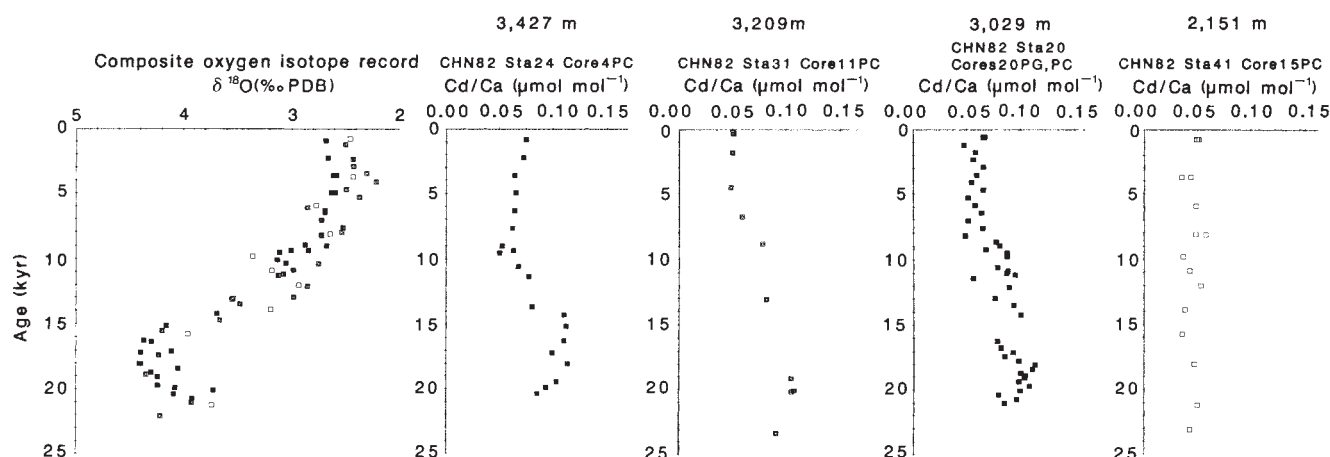


Fig. 2 Oxygen isotope and Cd data from benthic foraminifera from cores comprising a depth transect between 2,100 and 3,500 m at 42° N in the western North Atlantic ocean. The timescale (consistent with that used for EN120 GGC1) is chosen to produce the best overlap of the oxygen isotope records. The $\delta^{18}\text{O}$ plot indicates the relative success of the overlap. As the cores used here have lower sedimentation rates (varying from 1.5 to 5 cm kyr⁻¹), the resolution is poorer than for EN120 and should be considered reliable to no better than a few thousand years.

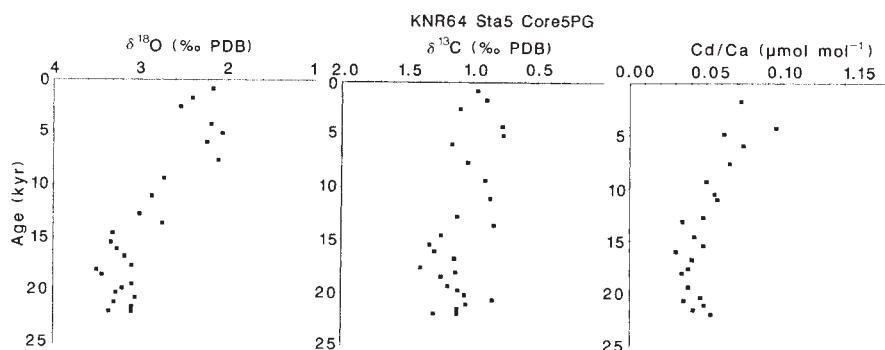


Fig. 3 Isotope and Cd data from benthic foraminifera from Caribbean Sea core KNR64 Sta5 Core5PG (16°31.3' N, 74°48.4' W, 3,047 m) plotted against estimated age. The age scale is derived from assuming that the physical top of the core (6 cm relative to the liner) is 0 kyr, that 16 cm is 8.5 kyr, that 22.5 cm is 14 kyr and that 41 cm is 18 kyr. The sedimentation rate of the core limits true resolution to ~5,000 years in the Holocene, somewhat better in the glacial. A large fraction of the total glacial/interglacial amplitude for $\delta^{18}\text{O}$ is observed, so that at least 70% of the Cd amplitude must be recorded.

glacial times. Furthermore, data from a Caribbean Sea core (KNR64 Sta5 Core5PC, 16°31.3' N, 74°48.4' W, 3,047 m) demonstrate that Caribbean nutrient concentrations decreased during the last glacial maximum (Table 2, Fig. 3). Reduced glacial Caribbean nutrients must be due to lower nutrient concentrations in source waters for the basin. The Caribbean derives its waters from ~1,700–1,800 m depth in the North Atlantic (the sill depth is 1,800 m, but some shallower water mixes in during inflow)²⁵. Ventilation of this basin is sufficiently rapid at present to minimize nutrient enhancements due to decomposing particles from the ocean surface. Modern source waters have phosphorus contents of ~1.5 $\mu\text{mol kg}^{-1}$ (ref. 19); phosphorus contents in the Caribbean are comparable, never being more than 0.2–0.4 $\mu\text{mol kg}^{-1}$ higher²⁶. Because the modern local nutrient input is small, significant reductions in Caribbean nutrient levels can be generated only by decreased nutrient concentrations in the advective sources. The decrease in glacial Caribbean nutrients (inferred from decreased Cd and higher $\delta^{13}\text{C}$) shows that Atlantic intermediate-depth waters were more nutrient-depleted at a time when deeper waters were more nutrient-enriched. These results can be expressed as a reconstructed profile of phosphorus in the glacial ocean (Fig. 4).

CLIMAP²⁷ included data showing more positive glacial Caribbean $\delta^{13}\text{C}$ in isotope stage 6 relative to stage 5e (although they did not discuss this). Cofer-Shabica and Peterson (in preparation) obtained $\delta^{13}\text{C}$ data from a long Caribbean core showing that glacial nutrient depletion occurs repeatedly during the late Quaternary (although they suggested that reductions in nutrient fluxes from the surface were responsible, which is not feasible). Higher intermediate-depth Atlantic $\delta^{13}\text{C}$ values were also reported for the glacial period preceding the last interglacial period ~140,000 yr ago²⁸ (although some caution should be applied in considering that data, because some of the

foraminifera used are now known to be unreliable²⁹). Subsequent to the present measurements and these other studies, Oppo and Fairbanks¹⁸ made measurements on two Caribbean cores reinforcing these results. Hence it is well established that glacial Caribbean and intermediate-depth Atlantic waters were nutrient-depleted. The Cd data indicate that the nutrient content decreased by a factor of two, equivalent to a phosphorus decrease from 1.6 to 0.8 $\mu\text{mol kg}^{-1}$.

It has not yet been established that the intermediate-depth ocean responded in the same way during the Younger Dryas cold event as it did during the glacial maximum. This may be due to the lack of a suitable high-sedimentation-rate core which could document this event. If there is a mechanistic link between changes in the deep and intermediate ocean, then we would expect to find such an event when suitable cores are recovered.

Possible mechanistic link

The model we propose derives from analogy with the modern North Pacific Ocean, where active intermediate-water production occurs, but no deep water is formed. Deep North Pacific waters can be traced only to southern deep-water sources³⁰. The immediate reason for this difference between the oceans is that North Pacific surface waters are relatively fresh (32.9%). Even when cooled to their freezing point, these waters cannot attain densities as high as those of Antarctic and North Atlantic deep and bottom waters. On winter cooling, these waters become sufficiently dense only to sink to intermediate levels. For this reason, a tongue of high-oxygen, low-salinity water is observed to spread from this source³¹. The low salinity of North Pacific surface waters is maintained by the balance between advection, evaporation, precipitation and runoff. Warren³⁰ has shown that precipitation is similar in the North Atlantic and North Pacific, but evaporation is lower in the North Pacific due to lower surface

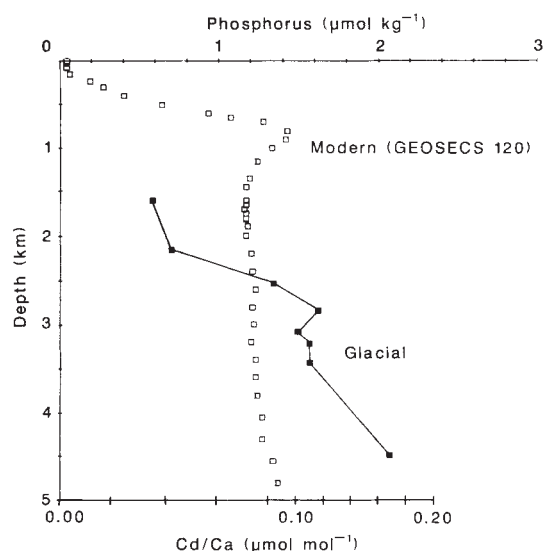


Fig. 4 Glacial depth profile of phosphorus in the modern western North Atlantic (GEOSECS station 120) compared to glacial profile. The glacial profile is based on the cores shown in Figs 1–3 as supplemented by two others whose glacial section is determined from percentage CaCO_3 curves (CHN82 Sta30-1 Core9PC, $41^\circ 50.6' \text{ N}$, $26^\circ 27' \text{ W}$, 2,830 m, glacial sample at 30–40 cm; CHN82 Sta23 Core3PC, $41^\circ 38' \text{ N}$, $27^\circ 20' \text{ W}$, 2,525 m; glacial sample at 36–42 cm). The actual data for the glacial profile are Cd/Ca ratios in benthic foraminifera (bottom scale). The phosphorus scale is estimated assuming that the Cd–P relationship in the glacial ocean is similar to that of today, and assuming that the distribution coefficient for Cd/Ca in foraminifera relative to sea water is 2.9. Note the change in the Cd scale, which is based on observation of a significant 'kink' in the Cd–P relationship (ref. 23). Actual glacial P is uncertain to the extent that this relationship may have been slightly different during glacial times and to the extent that the distribution coefficient is uncertain (it is known to no better than 15%). Despite these uncertainties, the important feature is the relative vertical homogeneity in the modern North Atlantic compared to the marked depth gradient in the glacial ocean. This qualitative difference depends neither on the assumed Cd–P relationship nor on the assumed distribution coefficient.

temperatures. Essentially, water 'distills' off the warm North Atlantic and is transported by the atmosphere to 'condense' onto the cool North Pacific⁶.

Proxy climate data show that the glacial high-latitude North Atlantic was much colder than at present^{32,33}. An analysis similar to Warren's³⁰ for the modern and glacial North Atlantic from 40 to 65° N and from 10° E to 70° W indicates the consequences of that cooling (Table 3). Neglecting the effect of changes in the wind speed and relative humidity over the glacial North Atlantic, cooler glacial temperatures would tend to reduce the rate of evaporation by a factor of two. This difference is comparable to the contrast between the modern North Atlantic and Pacific. This effect is reinforced by diminished northward transport of saline water from lower latitudes. Processes tending towards lower surface salinities are countered in part by increased wind strengths which enhance evaporation³⁴. Furthermore, dry catabatic winds cascading off continental ice masses would have reduced atmospheric relative humidity near the ice edge and contributed to the formation of cold saline water (J. Reid, personal communication). Hence, although it is plausible on the basis of temperature changes that the salinity of the Atlantic north of 45° N tended to lower values during glacial times (which has already been surmised qualitatively³²), it probably did not attain the state of the modern North Pacific.

This decrease in salinity would have made it more difficult to form very dense waters in the North Atlantic. It would have favoured more production of intermediate waters. The palaeochemical evidence shows that the glacial North Atlantic

Table 2 Data from two cores

CHN82 Sta 41 Core15PC			
Depth	Cd/Ca	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$
cm	($\mu\text{mol mol}^{-1}$)	(‰ PDB)	-
2	0.051	2.46	0.63
6	0.040	2.43	1.28
9	0.049	2.77	1.24
12	0.054	2.64	1.31
16	0.037	3.36	1.21
19	0.043	3.18	1.27
22	0.053	2.93	0.87
25		3.55	0.88
27	0.038	3.19	1.02
29	0.036	3.96	1.17
31	0.047	3.92	1.19
33		3.95	1.27
36	0.050	3.74	1.21
39	0.043		
41	0.048	3.42	1.18

KNR65 Sta5 Core5PG

7		2.17	0.96
8	0.073	2.41	0.89
9		2.54	1.09
11	0.095	2.19	0.77
12	0.061	2.05	0.76
13	0.074	2.24	1.15
15	0.065	2.11	1.03
17	0.050	2.73	0.90
18	0.055		
19	0.057	2.87	0.86
21	0.048	3.01	1.11
22	0.035	2.75	0.83
25	0.042	3.32	1.23
29	0.048	3.34	1.32
32	0.030	3.27	1.28
35	0.041	3.18	1.13
39	0.038	3.10	1.39
41	0.034	3.50	1.12
43		3.44	1.23
47	0.038	3.10	1.18
49		3.21	1.10
51		3.28	1.05
52	0.046		
53	0.035	3.06	0.84
55		3.30	1.04
55	0.048		
57	0.041	3.10	1.11
59	0.053	3.23	1.20

Cd data from *C. kullenbergi* and *C. wuellerstorfi*. Isotope data from *C. wuellerstorfi*. The time interval is verified as isotope stages 1 and 2 by the presence of excess ^{230}Th (ref. 38) and the absence of *G. flexuosa* (ref. 39).

did not fully attain the conditions of the modern North Pacific; some deep water must have been forming to maintain the Atlantic at lower nutrient levels than the Pacific^{16,24}. The North Atlantic probably shifted towards a state more like the North Pacific, with less deep-water production and more intermediate-water production. This model predicts that glacial North Atlantic intermediate waters would have been colder and fresher (and perhaps relatively more depleted in ^{18}O).

An alternative model

Another intermediate-water source for the modern Atlantic is the outflow of nutrient-depleted saline water from the Mediterranean. Zahn, Sarnthein and Erlenkeuser³⁵ studied the carbon isotope composition of sediments from intermediate depths in the eastern Atlantic from Gibraltar southwards. They found evidence for enhanced nutrient depletion of these mid-depth waters during glacial times. This evidence was interpreted as indicating an increased flux of nutrient-depleted Mediterranean

outflow water. Oppo and Fairbanks¹⁸ followed up this hypothesis to argue that the glacial intermediate-water nutrient depletion discussed above was due to this increased flux of outflow water from the Mediterranean. This enhanced Mediterranean outflow water/intermediate water hypothesis predicts that glacial intermediate waters were warmer and more saline, in contrast to the predictions of the glacial North Atlantic Intermediate Water hypothesis presented above. Observations that could establish the salinity or other characteristic properties of these water masses would be of help in distinguishing between these hypotheses.

Current evidence is not sufficient to disprove the Mediterranean outflow water hypothesis, but an unusual conjunction of circumstances is required by this model. Water from the Mediterranean contributes a very small fraction (<10%) of mid-depth water in the modern Atlantic. All other sources of intermediate water would have to wane before Mediterranean outflow could become the dominant source and determinant of intermediate-water nutrients. In view of the favourable conditions for North Atlantic Intermediate Water formation discussed above, a substantial reduction in all other sources seems unlikely. Furthermore, the degree to which Mediterranean water outflow could increase is constrained by the relationship between evaporation, density contrast and hydraulic controls, which suggests that there should not be major changes in the rate of outflow (ref. 36; R. C. Thunell, personal communication). It is even possible that the mid-depth nutrient depletion observed in the eastern Atlantic³⁵ is simply due to impingement of glacial North Atlantic Intermediate Water on the continental margin, and that no enhancement of Mediterranean outflow is necessary.

Conclusions and speculations

A link between the ocean surface and deep waters is now established for one of the most rapid extreme climate events. On timescales of several hundreds to several hundred thousand years, colder surface temperatures in the high-latitude North Atlantic are associated with reduced fluxes of nutrient-depleted Northern source water into the deep Atlantic. During the last glacial maximum, this association is accompanied by intermediate-water nutrient depletion. We propose that lower glacial North Atlantic surface temperatures reduce evaporation and hence decrease high-latitude surface salinities. The lower potential density of these surface waters when cooled leads to the formation of glacial North Atlantic Intermediate Water at the expense of reduced formation of glacial North Atlantic Deep Water.

This hypothesis might be taken to imply that the surface ocean controls a passive deep ocean. To the extent that intermediate- and deep-water events are coupled, this implication is intended. But it is not clear whether surface ocean temperatures control

Table 3 Evaporation in the modern and glacial North Atlantic

	February	August	Annual average
Modern	26.6	12.7	19.7
Glacial	15.1	8.3	11.7

The calculations were carried out as described by Warren⁶. The block size was 5 degrees of latitude by 10 degrees of longitude. The results are expressed as the areal average of the product of wind speed times the specific humidity difference between air and saturation at the sea surface temperature (units: mm s⁻¹). To convert these values to evaporation rates in mm s⁻¹, it is necessary to multiply by the density of air (1.25×10^{-3} g cm⁻³) and by the exchange coefficient for moisture, which has been estimated in the range $(1.2-2.1) \times 10^{-3}$ (ref. 6). The area of open water is slightly less during the glacial period, but it is not clear exactly how sea-ice cover should be treated for the evaporation-precipitation cycle.

deep-water formation, or the reverse: warm surface temperatures in the modern North Atlantic may be due in part to the requirement for northward flow of warm surface water to compensate southward flow of cold deep water. The North Atlantic Current is wind driven and hence controlled by the atmospheric circulation. But on a global scale, atmospheric wind patterns are also affected in turn by the distribution of ocean surface temperatures. Held³⁷ suggests that orographic controls on atmospheric circulation (the relative positions of the Rockies and Himalayas) may account in part for differences in wind stress over the North Atlantic and North Pacific. If this is so, then the primary cause of warm North Atlantic surface temperatures is the topography of the continents coupled to atmospheric dynamics, with the ocean playing a relatively passive role. The presence of a large ice mass on the North American continent has a dramatic orographic effect on atmospheric circulation patterns. Modelling studies³⁴ suggest that these changes could account for the southward movement of the polar front during the last glacial maximum. The response of the planetary wave system to changes in position and elevation of the wasting glacial ice mass could possibly have led to the Younger Dryas cooling. Other plausible causes have been suggested^{6,7}, so this remains speculative. In any event, such considerations illustrate how the long-term progression of glaciation might be linked to both orbitally induced insolation redistribution and the dynamics of the fluid and solid Earth.

We thank Carl Wunsch and Joe Reid for comments which have been useful in developing the hypothesis outlined here. We thank John Jasper, Jim Broda and Ed Laine for advice and use of their coring equipment, the officers and crew of the RV *Endeavor*, Chris Measures and Tim Jickells for their assistance in obtaining the core, and C. Eben Franks for isotopic analyses. This research was supported by the NSF, as is curation of cores.

Received 21 May; accepted 16 September 1987.

- Broecker, W. S. in *Milankovitch and Climate* (eds Berger, A. et al.) 687-698 (Reidel, Dordrecht, 1984).
- Duplessy, J.-C., Delibrias, G., Turon, G., Pujol, C. & Duprat, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 121-144 (1981).
- Mix, A. C. & Ruddiman, W. F. *Quat. Sci. Rev.* **4**, 59-108 (1985).
- Bloom, A. L. in *Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 355-380 (Yale University Press, 1971).
- Lowe, J. J., Gray, J. M. & Robinson, J. E. *Late Glacial of Northwest Europe* (Pergamon, Oxford, 1980).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21-26 (1985).
- Broecker, W. S. et al. *Palaeoceanography* (in the press).
- Dansgaard, W. *Am. geophys. Un. Monogr.* **29**, 288-298 (1984).
- Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284-291 (1986).
- Rind, D., Peteet, D., Broecker, W., McIntyre, A. & Ruddiman, W. *Clim. Dyn.* **1**, 3-33 (1986).
- Keigwin, L. D., Corliss, B. H., Druffel, E. R. M. & Laine, E. P. *Quat. Res.* **22**, 383-386 (1984).
- Laine, E. P. & Hollister, C. D. *Mar. Geol.* **39**, 277-300 (1981).
- Shackleton, N. J. & Chappell, J. *Nature* **324**, 137-140 (1986).
- Broecker, W. S. & Peng, T. H. *Tracers in the Sea* (Eldigio, Palisades, 1982).
- Duplessy, J.-C. et al. *Nature* **320**, 350-352 (1986).
- Boyle, E. A. & Keigwin, L. D. *Earth planet. Sci. Lett.* **76**, 135-150 (1985).
- Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784-787 (1982).
- Oppo, D. & Fairbanks, R. G. *Earth planet. Sci. Lett.* (in the press).
- Bainbridge, A. E. *GEOSecs Atlantic Ocean Atlas* (US Government Printing Office, Washington, 1980).
- Craig, H., Broecker, W. S. & Spencer, D. W. *GEOSecs Pacific Ocean Atlas* (US Government Printing Office, Washington, 1982).
- Weiss, R., Broecker, W. S., Craig, H. & Spencer, D. W. *GEOSecs Indian Ocean Atlas* (US Government Printing Office, Washington, 1982).
- Boyle, E. A. *Geochim. cosmochim. Acta* **50**, 265-276 (1985).
- Boyle, E. A. *Palaeoceanography* (in the press).
- Boyle, E. A. *Am. geophys. Un. Monogr.* **29**, 360-368 (1984).
- Stalcup, M. C., Metcalf, W. G. & Johnson, R. G. *J. mar. Res. Suppl.* **33**, 15-35 (1975).
- Richards, F. A. *J. mar. Res.* **17**, 449-465 (1958).
- CLIMAP project members *Quat. Res.* **21**, 123-124 (1984).
- Duplessy, J.-C. & Shackleton, N. J. *Nature* **316**, 500-507 (1985).
- Zahn, R., Winn, K. & Sarthein, M. *Palaeoceanography* **1**, 27-42 (1986).
- Warren, B. A. *J. mar. Res.* **41**, 327-347 (1983).
- Reid, J. L. *Johns Hopkins Oceanographic Studies no. 2* (Johns Hopkins Press, Baltimore, 1965).
- McIntyre, A. et al. *Mem. geol. Soc. Am.* **145**, 43-76 (1976).
- CLIMAP project members *Geol. Soc. Am. Map and Chart Ser. No. 36* (1981).
- Manabe, S. & Broccoli, A. J. *J. geophys. Res.* **90**, 2167-2190 (1985).
- Zahn, R., Sarthein, M. & Erlenkeuser, M. *Palaeoceanography* (in the press).
- Bryden, H. & Stommel, H. *Oceanol. Acta* **7**, 289-296 (1984).
- Held, I. M. in *Circulation in the Atmosphere* (eds Hoskins, B. & Pierce, B.) (Academic, New York, 1982).
- Ku, T.-L., Bischoff, J. L. & Boersma, A. *Deep-Sea Res.* **19**, 233-247 (1972).
- Erickson, D. B. in *The Sea Vol. 1* (ed. Hill, M. N.) (Interscience, New York, 1963).

Contributions of hydrogen bonds of Thr 157 to the thermodynamic stability of phage T4 lysozyme

Tom Alber*, Sun Dao-pin, Keith Wilson, Joan A. Wozniak, Sean P. Cook† & Brian W. Matthews

Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

Measurements of changes in structure and stability caused by 13 different substitutions for threonine 157 in phage T4 lysozyme show that the most stable lysozyme variants contain hydrogen bonds analogous to those in the wild-type enzyme and that structural adjustments allow the protein to be surprisingly tolerant of amino-acid substitutions.

ALTHOUGH the three-dimensional structure of a protein can be determined relatively accurately, the noncovalent interactions that stabilize a particular conformation have escaped quantitative enumeration. This gap between understanding structure and energy has hampered the design and prediction of protein structures from amino-acid sequence data. It has also made it difficult to predict the consequences of chemical changes in a protein sequence. Major barriers are presented by the complexity of the folded and unfolded states and their interactions with solvent. The thermodynamics of denaturation have also proved to be complicated. Proteins have a temperature of maximum stability, and denaturation induced by heat and by cold have been characterized¹⁻³.

A traditional approach to determining the contribution of specific interactions to the thermodynamic stability of a protein is to compare the properties of variants with unique chemical modifications or amino-acid substitutions. Changes in stability are ascribed to the site that is chemically or genetically altered. This approach has been used to probe the structural basis of the thermodynamic stability of phage T4 lysozyme^{2,4-7}. Temperature sensitive (ts) mutations experimentally identify essential interactions^{8,9}. The relationships between the structure, dynamics and stability of the wild-type and mutant proteins have been investigated using X-ray crystallography^{9,10}, fluorescence¹¹, nuclear magnetic resonance^{12,13} and thermodynamic methods^{2,4-7}.

Here we have used oligonucleotide-directed mutagenesis to replace residue Thr 157 in phage T4 lysozyme with 13 other amino acids. The mutant proteins were purified and subjected to thermodynamic measurements of stability and high resolution X-ray structural analysis. Comparisons of the structures and stabilities of the mutant and wild-type phage lysozymes reveal the relative contributions of each substituent of the Thr 157 side chain to the overall stability of the protein.

Thr 157 substituted by Ile

Thr 157 is located in an irregular loop on the surface of the protein¹⁴. The protein having of this residue substituted by Ile was previously isolated as a temperature sensitive (ts) mutant in a screen of randomly mutagenized T4 phage¹⁰. The mutation reduces the stability of the purified protein by about 2.9 kcal·mol⁻¹ at 42°C, pH 2.0. The refined X-ray crystal structure of the mutant protein showed that the replacement of the hydroxyl group of Thr with an ethyl group (Ile) leads to a complex, though localized, structural change¹⁰ (Fig. 1a). Hydrogen bonds to the Thr 157 hydroxyl group are lost in the mutant protein, but tighter binding of a water molecule in the vicinity of the Thr 155 hydroxyl group may partially compensate for this

change. The β -carbon of residue 157 shifts ~ 0.7 Å, and a rotation about the C_{α} - C_{β} bond causes a change in the contacts of the side chain γ -methyl group. The side chain of Asp 159 moves ~ 1.1 Å to accommodate the bulkier ethyl group at position 157.

The main reason for the destabilizing effect of the Thr 157 to Ile substitution appears to be the lack of a hydrogen bond acceptor for the buried main chain amide of Asp 159 in the mutant protein¹⁰. But the observed structural shifts could also materially alter van der Waals contacts, solvation, torsional interactions, hydrophobic forces and electrostatic interactions. The complexity of this result illustrates a difficulty with the genetic approach to dissecting protein stability; even simple changes can alter several interactions in the folded state. The inability of thermodynamic methods to distinguish between the effects of a mutation on the folded and unfolded states presents a further complication.

Mutant lysozymes

The common amino acids provide good probes of the contributions of each of the substituents of the threonine side chain. Serine provides a measure of the contribution of hydrogen bonds as it contains a stereochemically equivalent side-chain hydroxyl group. Valine tests the importance of van der Waals and hydrophobic forces because it is similar in shape to Thr but cannot hydrogen bond. Alanine contains only a β -methyl group and glycine lacks a side chain altogether. Asparagine may provide new polar interactions, and charged residues can be introduced to probe electrostatic effects. Nonpolar amino acids can help to determine if stability is correlated with hydrophobicity rather than with particular structural features of the native state.

Oligonucleotide-directed mutagenesis was used to construct lysozymes with Ser, Val, Ala, Gly, Asn, Cys, His, Arg, Asp, Glu,

Table 1 Mutagenic oligonucleotides

Oligonucleotide mixture	Amino acids encoded at position 157
CC	
5' GCGTCCCAATTGCCAGTT 3'	Val, Leu, Ile, Ser, Gly, Arg, His, Asp, Asn
AG	
-----ACC-----	Val, Gly, Phe, Cys
AA	
-----ANC-----	Val, Ala, Asp, Glu, Gly
-----NAC-----	Val

Primers were made using DNA synthesizers (Biosearch and Applied Biosystems) and 18-mers were purified by electrophoresis through 8 M urea-15% polyacrylamide gels¹⁷. The oligonucleotides specifying Val 157 migrated anomalously fast in such gels. As a result, unpurified oligonucleotide mixtures have been used in subsequent experiments using this method. N is A, G, T and C.

Present addresses: *Department of Biochemistry, University of Utah School of Medicine, 50 N Medical Drive, Salt Lake City, Utah 84132, USA. † Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA.

Leu, Phe and Ile at position 157. Mixtures of mutagenic 18 mers (Table 1) were used in the two-primer protocol of Zoller and Smith¹⁵ to produce different mutations rapidly¹⁶. To produce the altered proteins, the 3' half of each mutant lysozyme gene was joined to wild-type 5' coding sequences contained on the expression vector pHSe5 (amp^r, tet^r, laqI^r, Δ T4e) (ref. 16). The proteins were purified by chromatography on CM-Sephadex¹³ and shown to be homogeneous by SDS-gel electrophoresis^{18,19} and reverse-phase high-pressure liquid chromatography²⁰ (not shown).

Thermodynamic studies

The thermodynamic stability of the mutant lysozymes was assessed by measuring the molar ellipticity of a solution of the purified protein as a function of temperature⁶. The wavelength used (223 nm) is sensitive to the presence of secondary structure in the protein. As T4 lysozyme unfolds, the optical signal undergoes a cooperative transition. Assuming that only the folded and denatured states contribute to the measured molar ellipticity, the low- and high-temperature baselines were extrapolated into the transition region and the fraction of material that is folded was calculated as a function of temperature. At the midpoint temperature of the denaturation transition, T_m , half the protein is unfolded, the equilibrium constant for the unfolding reaction is unity and ΔG is zero. The stability of each mutant protein at the T_m of wild-type lysozyme was estimated using the relation $\Delta G = \Delta T_m \Delta S$ (ref. 7). The change in T_m caused by each substitution at position 157 and the corresponding change in the free energy of stabilization is shown in Table 2.

The stability of the proteins listed in Table 2 is not obviously correlated with the charge, size or hydrophobicity of the side chain at position 157. Side chains with polar atoms do seem to make a larger contribution to stability than strictly nonpolar side chains. Although this supports the view that hydrogen bonds to this residue are critical, there are apparent exceptions. Gly, for example, lacks a side chain altogether, but provides almost as much stability as Ser. His is one of the most poorly accommodated amino acids, but similar as in structure to Asn, the most stabilizing mutant residue.

X-ray structural studies

To investigate the structural basis of the observed stabilities, the X-ray crystal structure of each mutant lysozyme was determined at high resolution. Isomorphous crystals of the mutant and wild-type enzymes were grown using phosphate as the precipitant²². Crystals were obtained in the pH range of 6.4–7.1. Although residue 157 is not near an intermolecular contact, some mutant lysozymes, notably Thr 157 substituted by Arg, rapidly formed bigger crystals than the wild-type protein. Other variants were more difficult to crystallize. Before data collection, the crystals were soaked in a standard mother liquor containing 1.05M K_2HPO_4 , 1.26M Na_2HPO_4 , 0.23 M NaCl, 1.4 mM β -mercaptoethanol, pH 6.7.

High-resolution X-ray data were collected by oscillation photography²³ (Table 3). After data processing²⁴, electron density maps were calculated using coefficients ($F_{\text{mutant}} - F_{\text{wild type}}$) (Fig. 1b) and ($2F_{\text{mutant}} - F_{\text{wild type}}$) (Fig. 1c) and the phases of the refined structure of the wild-type protein¹⁴. These maps were used to build starting models of the mutant proteins, using the molecular graphics program FRODO (ref. 25). Restrained least-squares refinement of the atomic coordinates against the X-ray data was carried out with the TNT package of programs²⁶. Data processing and refinement statistics are given in Table 3.

The different side chains introduced at position 157 display a variety of conformations. As in the case of the original ts mutant (Thr 157 substituted by Ile), the X-ray crystal structures of the engineered variants show that all the substitutions are accommodated with only localized changes in the protein and solvent. The smallest structural change is induced by Ser. In

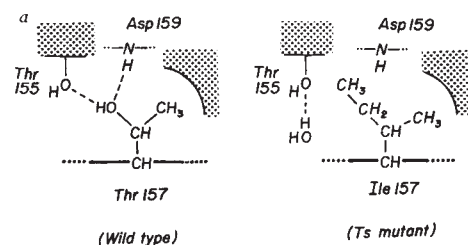


Fig. 1 *a*, Schematic illustration comparing the local environment of Thr 157 in wild-type lysozyme and Ile 157 in the ts mutant structure. Hydrogen bonds are indicated by broken lines. *b*, Representative map showing the difference in density between a mutant (His 157) lysozyme and the wild-type (Thr 157) enzyme. Amplitudes ($F_{\text{His 157}} - F_{\text{Thr 157}}$); phases calculated from the refined model of wild-type lysozyme¹⁴; resolution 1.9 Å. Positive contours (blue) and negative contours (orange) are drawn at levels of $\pm 4\sigma$, where σ is the root-mean-square density throughout the unit cell. The large negative feature corresponds to the loss of the threonine γ -hydroxyl group in the His 157 structure. The large positive feature shows the location of the introduced imidazole ring. There is no density at the β -carbon and γ -carbon positions, which are common to the His 157 and Thr 157 structures. *c*, Representative map showing the electron density of the mutant lysozyme with histidine at position 157. The imidazole ring extends into the solvent. The two bound solvent molecules (OH256 and OH312) are also seen in the wild-type structure. Coefficients ($2F_{\text{His 157}} - F_{\text{Thr 157}}$), phases from ref. 14, resolution 1.9 Å. *d*, Electron density of the mutant lysozyme with glycine at position 157 superimposed on the refined model of the wild-type protein. In the Gly 157 structure a water molecule binds close to the site previously occupied by the γ -hydroxyl of Thr 157 and maintains the hydrogen-bond network to the amide of Asp 159 and the hydroxyl of Thr 155. Amplitudes ($2F_{\text{Gly 157}} - F_{\text{Thr 157}}$), phases from ref. 14, resolution 1.9 Å. *e*, Superposition of the refined side-chains of Thr 157 (yellow) and Asn 157 (blue) showing the near-superposition of the γ -hydroxyl of the threonine and the δ -oxygen of the asparagine.

this case, the coordinates of equivalent atoms in the mutant and wild-type proteins superimpose within the experimental error (data not shown). In particular, the hydroxyl group of Ser 157 occupies the same position and participates in the same interactions as the hydroxyl group of Thr 157.

In terms of protein stability, Asn, Ser, Asp and Gly are the best substitutes for Thr (Table 2). The X-ray crystal structures of these mutant lysozymes support the importance of hydrogen bonding at this site. To accommodate Asn, the main chain and the β -carbon shift to allow the longer polar side chain to come close to the groups that interact with the Thr 157 hydroxyl group in the wild-type protein (Fig. 1d, Fig. 3 and Table 4). Asn and Thr can donate different numbers of protons to hydrogen bonds so the pattern of hydrogen bonds is likely to differ in detail in the mutant and wild-type lysozymes. Nonetheless, the Asn 157 mutant is almost as stable as the wild-type protein (Table 2).

Asp is also one of the most stabilizing substitutions for Thr 157. In contrast to Asn 157, Asp is best modelled in two distinct positions, each occurring for about half of the time. One conformer resembles the structure of the Asn 157 mutant. One oxygen of the side chain carboxylate is within hydrogen-bonding distance of the hydroxyl group of Thr 155, the main chain amide of Asp 159 and the side chain of Asp 159 (Fig. 3, Table 4). In the second conformer, an $\sim 120^\circ$ rotation of χ_1 brings the Asp 157 side chain into a torsionally favourable orientation on the opposite side of the 157–158 peptide plane. This results in close contact (2.4 Å) between the 157 side chain carboxylate oxygen and the main chain amide of Trp 158. This conformer displaces a water molecule that binds to the 158 amide in the wild type and all other mutant proteins. This alternate conformer may arise from electrostatic repulsion between Asp 157 and Asp 159. The side chain rotation increases the distance between these carboxylates from 3.2 Å to 4.9 Å.

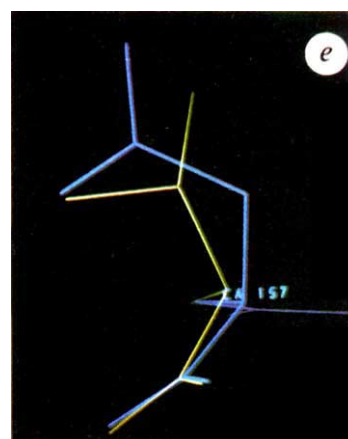
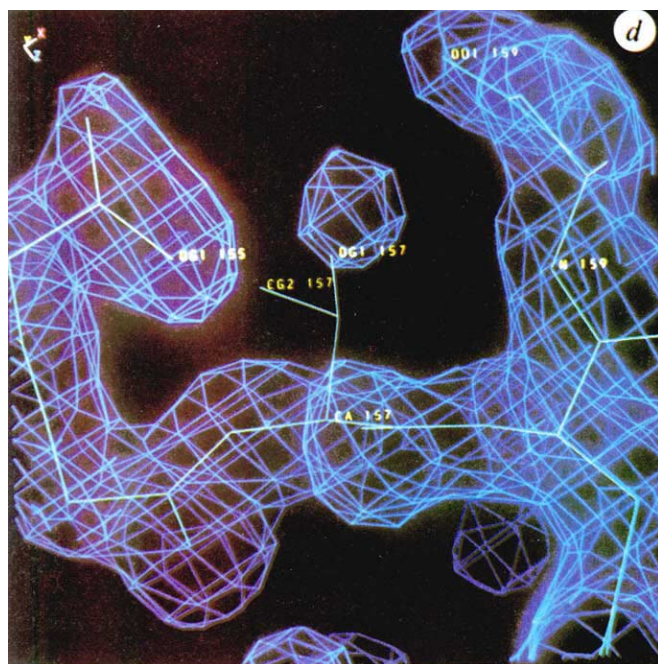
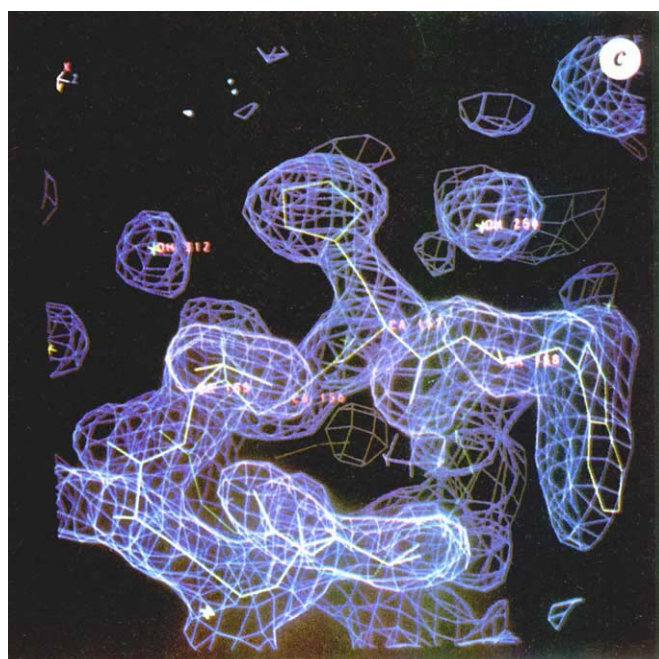
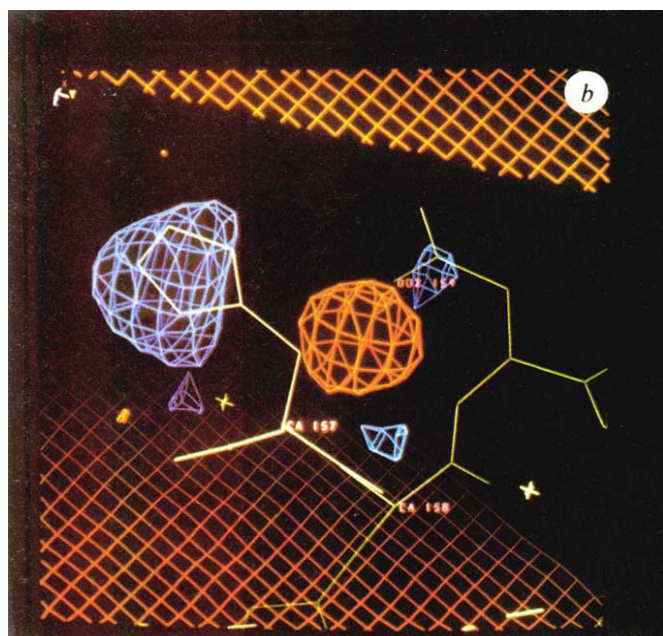
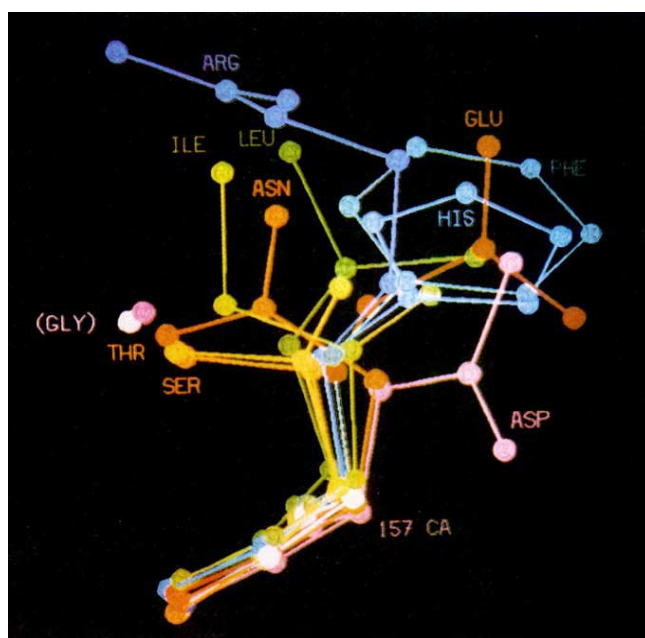


Fig. 2 (Right) Superposition of residue 157 in wild-type and mutant lysozymes. The name of each amino acid has the same colour as the residue itself. The oxygen atoms of Thr, Ser and Asn that participate in similar hydrogen-bond networks (see text) are clustered at the left of the figure. Also included in this cluster is a water molecule (white) bound to Gly 157 lysozyme. The side-chain of Asp 157 adopts two alternative conformations; in the case shown a water molecule (pink) binds at the 'Gly' position. The second conformation (not shown), is very similar to that of Asn 157. The configuration of Val 157 is superimposable on that of Ile 157 and is not included in the figure.



Perhaps the most striking X-ray crystal structure of this series is that of the mutant protein with Gly at position 157 (Fig. 1d). The absence of a side-chain leaves a gap on the surface of the enzyme that is filled by a new bound water molecule. The thermal factor of this solvent molecule ($29 \text{ \AA}^2$) is well below the average for ordered solvent in this structure ($\sim 48 \text{ \AA}^2$). The water molecule is positioned to interact with the same hydrogen bonded partners as Thr 157 in the wild-type protein. The average geometry of the hydrogen bond network is improved, apparently because the solvent molecule is not covalently bonded to the protein. With two protons to donate, the new water can also form a hydrogen bond with the side chain carboxylate of Asp 159 (Fig. 3 and Table 4). In the wild-type protein, the stereochemistry of the hydrogen bond network in this region argues that Thr 157 and Asp 159 can form a hydrogen bond only at the expense of a hydrogen bond between the hydroxyl group of Thr 155 and the buried carbonyl oxygen of Thr 151 (ref. 10).

At least two factors may mitigate the contribution of the improved hydrogen bonding geometry to the stability of the Gly 157 mutant protein. First, the binding of the water molecule is an intermolecular reaction with a substantial loss of entropy expected. In addition, the new glycine in the polypeptide chain may stabilize the unfolded state by increasing the number of accessible conformations^{27,28}. Nonetheless, the X-ray crystal structure of the Gly 157 mutant demonstrates that hydrogen bonding is sufficient to confer the bulk of the free energy of stabilization provided by the Thr 157 side chain. It is also apparent that solvent binding to a specific site in the native state can confer protein stability.

Binding of the new water molecule observed in the Gly 157 mutant is blocked by any side chain at this position that includes a β -carbon. The β -methyl group of Ala 157, for example, is approximately $2.5 \text{ \AA}$ from the water binding site. As a result, the buried main chain amide of Asp 159 is not paired with a hydrogen bond acceptor in the X-ray crystal structures of the proteins with Ala, Arg, Val, Leu, Glu, His, Phe or Ile at position 157. These mutant lysozymes are all less stable than the proteins in which the amide of Asp 159 participates in a hydrogen bond (Table 2).

The X-ray crystal structures of the mutants which lack a hydrogen bond to the amide of Asp 159 fall into two distinct classes. The first includes lysozymes with Val and Ile at position 157. The stereochemically equivalent atoms of these β -branched side chains occupy indistinguishable positions in the X-ray crystal structures of the respective mutant proteins. The replacement of the Thr 157 hydroxyl group with an aliphatic carbon results in a rotation around the C_α - C_β bond. This partially

Table 2 Thermodynamic stability of lysozymes with different amino acids at position 157

Amino acid at position 157	ΔT_m (°C)	$\Delta\Delta G$ (kcal mol ⁻¹)
Thr (wild type)	—	—
Asp	-1.7	0.45
Ser	-2.5	0.66
Asp	-4.2	1.1
Gly	-4.2	1.1
Cys	-4.9	1.3
Leu	-5.0	1.3
Arg	-5.1	1.3
Ala	-5.4	1.4
Glu	-5.8	1.5
Val	-6.0	1.6
His	-7.9	2.1
Phe	-9.2	2.4
Ile (ts mutant)	-11.0	2.9

The proteins were purified as described¹³, except that cell lysis was promoted with EDTA (ref. 21) instead of chloroform. ΔT_m is the change in the melting temperature of the mutant lysozyme ($\pm 0.5^\circ$) relative to wild-type lysozyme ($T_m = 42^\circ\text{C}$ at pH 2.0). $\Delta\Delta G$ is the corresponding change in free energy at 42°C (ref. 7). Protein samples ($20\text{--}30 \mu\text{g ml}^{-1}$) were extensively dialysed against oxygen-free buffers. Ionic strength was 0.2 with KCl and pH was adjusted to 2.0 with HCl to avoid irreversible aggregation at high temperature. Circular dichroism ($\theta_{223} \text{ nm}$) was measured with a Jasco J-500C instrument equipped with a Hewlett Packard 89100A thermoionic controller. The temperature of the sample was changed at a constant rate of 1 K min^{-1} . To ensure reversibility, unfolding and refolding were both monitored^{5,6}.

relieves a close approach to the hydroxyl group of Thr 155 and alters van der Waals contacts to the γ -methyl group of residue 157. The Val and Ile 157 mutants clearly differ, however, in two ways. The solvation of Thr 155 is more pronounced with Ile at position 157 compared to Val. In addition, the bulkier ethyl substituent of Ile forces the side chain of Asp 159 to move $1.1 \text{ \AA}$ away from its position in the wild-type protein, compared to a movement of $0.4 \text{ \AA}$ in the Val 157 mutant protein. The unusually large motion of Asp 159 is correlated with a reduction in stability of $\sim 1 \text{ kcal mol}^{-1}$ at 42°C of the Ile 157 lysozyme compared to the mutant protein containing Val 157 (Table 2).

The second group of mutants lacking a hydrogen bond to the amide of Asp 159 includes the proteins with Ala, Leu, Glu, His, Phe and Arg at position 157. The β - and γ -methylene groups in these side chains follow the path of the β - and γ -carbons of

Table 3 Data processing and refinement statistics

Amino acid at position 157	Unit cell dimensions		Resolution ($\text{\AA}$)	Number of reflections	R_{merge} (%)	ΔBond ($\text{\AA}$)	ΔAngle ($^\circ$)	R (%)
	a, b ($\text{\AA}$)	c ($\text{\AA}$)						
Thr	61.2	96.8	1.7	17,021	14.3	0.018	2.3	19.3
Asp	61.2	97.1	1.7	15,983	9.3	0.019	2.6	17.6
Ser	61.2	97.1	1.7	16,232	6.7	0.019	2.4	17.6
Asp	61.2	97.2	1.7	15,179	6.0	0.021	2.8	18.1
Gly	61.2	97.1	1.7	16,167	7.8	0.022	2.7	18.7
Cys	61.2	96.8	1.7	14,463	7.9	0.018	2.6	16.4
Ala	61.1	97.3	1.7	15,193	7.8	0.020	2.7	18.3
Arg	61.1	97.2	1.7	16,037	6.0	0.020	2.6	17.4
Val	61.1	97.0	1.7	15,405	5.1	0.019	2.7	17.5
Leu	61.2	96.7	1.7	16,141	7.3	0.021	2.7	18.1
Glu	61.2	97.4	1.7	15,791	6.7	0.019	2.7	17.4
His	61.2	96.8	1.7	16,494	9.0	0.020	2.9	17.9
Phe	61.2	97.3	1.7	15,390	5.8	0.018	2.8	17.7
Ile	61.2	96.8	1.7	16,094	5.6	0.019	2.6	17.4

R_{merge} is the agreement between intensities on different films; ΔBond and ΔAngle are the root mean square deviations of the bond lengths and angles in the crystallographically refined structure from ideal stereochemistry²⁶. R is the crystallographic residual for the refined structure.

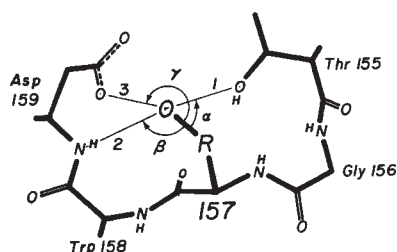


Fig. 3 Schematic diagram showing the geometry of the hydrogen-bond network formed with different amino acids at position 157. The three potential hydrogen bond lengths (1, 2, 3) and angles (α , β , γ) are listed in Table 4.

Thr 157 in the wild-type protein (Figs 1b, 2). This preserves the van der Waals contacts of the aliphatic portion of the Thr 157 side chain. These substitutions, however, are among the least stabilizing (Table 2). This result emphasizes the importance of the hydrogen bonds to the Thr 157 hydroxyl group for the stability of the wild-type protein.

The side chains of His 157 and Arg 157 make additional polar interactions. The conformation of His 157 is stabilized by hydrogen bonds to ordered solvent molecules (Fig. 1c). These water molecules are in similar positions in both the wild-type and His 157 mutant proteins. Arg 157 forms a new hydrogen bonded ion pair with Asp 159. The stability of the Arg 157 lysozyme and of the wild-type enzyme are more alike at pH 6.5 than at pH 2.0 (not shown). This suggests that the new ion-pair can partially compensate for the loss of stability arising from the buried unpaired amide of Asp 159 (to be discussed in detail elsewhere).

Even though torsional angles were not restrained during crystallographic refinement, the χ_1 angles of all the 157 side chains cluster around 3 values which differ by $\sim 120^\circ$ (Fig. 2). These values correspond to the minima in the torsional potentials of the isolated amino acids²⁹. Most of the side chain conformations at position 157 correspond to frequently observed rotomers³⁰. A notable exception is Asn 157, with $\chi_1 = 55^\circ$ and $\chi_2 = 53^\circ$. This conformation is found for only 6.6% of the asparagines in X-ray crystal structures of 19 proteins refined at high resolution³⁰. Nonetheless, Asn 157 is the best replacement for Thr. This suggests that rare side chain rotomers can form stabilizing interactions.

Implications for protein stability

All of the substitutions of Thr 157 produce lysozymes with stability between the wild-type and the original ts mutant protein. Stability is maximized by the wild-type Thr residue, but the protein is surprisingly tolerant of substitutions. In most cases, only small reductions in stability occur. The folded structure is seen to adjust to mitigate potentially deleterious effects of different substitutions.

The most severely destabilizing substitution is Thr 157 by Ile, originally isolated as a ts mutant¹⁰. Comparison of the X-ray crystal structures of the wild-type and ts mutant proteins showed differences in hydrogen bonding, hydration, van der Waals contacts and ionic interactions¹⁰, but the relative magnitudes of these effects could not be differentiated. Determination of the thermodynamic stabilities (Table 2) and the high resolution X-ray crystal structures (Fig. 2) of the 13 mutant proteins described here has revealed patterns of structure that correlate with stability.

The clearest pattern is that the buried main chain amide of Asp 159 is hydrogen-bonded in the most stable lysozyme variants (Thr, Asn, Ser, Asp and Gly 157). This correlation between structure and thermal stability suggests that the loss of this hydrogen bond in the folded protein dominates the effects of all thirteen substitutions on the free energies of the unfolded

Table 4 Geometry of the hydrogen-bond network formed with different amino acids at position 157

Amino acid	Atom	Interatomic distance ($\text{\AA}$)			Angle ($^\circ$)			B ($\text{\AA}^2$)
		1	2	3	α	β	γ	
Thr	O $^\gamma$	2.8	3.2	3.2	119	103	106	20
Ser	O $^\gamma$	2.8	3.4	3.2	119	107	114	30
Asn	O $^\delta$	3.1	3.1	3.2	97	130	101	30
Asp(1)	O $^\delta$	2.9	3.1	3.1	100	125	94	—
Gly	H $_2$ O	2.9	3.1	2.8	—	—	—	29

Interatomic distances 1, 2, 3 are the three potential hydrogen bond lengths and their angles are given by α , β , γ . Asp (1) is the conformation of aspartic acid 157 that resembles Asn 157 (see Fig. 2 legend). Also included is the crystallographic thermal factor, B , of the central hydrogen-bonded atom.

states of the mutant lysozymes. The correlation supports the previous conclusion that one of the most deleterious effects of the Thr 157 substitution by Ile is the loss of hydrogen bonds in the folded structure¹⁰. These results underscore the importance of pairing all hydrogen-bonding groups in the folded state³¹.

Nonetheless, it is apparent that different hydrogen bonds are associated with different net contributions to stability. Hydrogen bonds to the amide of Asp 159 are different in detail with Asn, Asp, Ser and Gly at position 157 (Fig. 3 and Table 4). These or other differences in the folded and the unfolded states of these variants could account for the observed small differences in stability. New hydrogen bonds between His 157 and bound solvent (Fig. 1c) or between Arg 157 and Asp 159 (not shown) do not fully compensate for the unpaired amide of Asp 159. Additional solvent binding to Thr 155 also does not rescue the ts phenotype. These results leave open the possibility that some hydrogen bonds can contribute excess free energy of stabilization to the folded state³¹⁻³⁴.

Hydrogen bonds to residue 157 are not sufficient to provide the observed stability of the wild-type enzyme. This is shown by the comparison of the proteins with Thr and Ser at position 157. The similarity of the structures of these lysozymes, taken together with the reduction in stability caused by Ser 157 (Table 2), indicate that the Thr γ -methyl group contributes ~ 0.66 kcal per mole to the stability of the wild-type protein. The contribution of the γ -methyl group could be direct (such as by increasing van der Waals contacts or burying hydrophobic surface area in the folded state) or indirect (for example, by decreasing the flexibility of the unfolded state^{27,28} or increasing the rigidity of hydrogen bonds to the γ -hydroxyl group³⁴). The last possibility is supported by the finding that the Ser 157 γ -hydroxyl group is less well-ordered ($B = 30 \text{ \AA}^2$) than the Thr 157 γ -hydroxyl ($B = 20 \text{ \AA}^2$), even though these groups occupy equivalent average positions.

The direct contribution of the Thr 157 γ -methyl group can be estimated from its static surface area ($17 \text{ \AA}^2$) (ref. 10) buried in the folded protein, but different methods lead to qualitatively different predictions for the function of the methyl group. On the basis of the free energies of transfer of hydrocarbons from organic solvents to water, scalar multipliers of 22–25 cal per $\text{\AA}^2$ have been proposed to estimate the free energy of stabilization from buried surface area³⁵. This approach suggests that the measured contribution of the methyl group ($0.66 \text{ kcal mol}^{-1}$) is largely accounted for ($0.43 \text{ kcal mol}^{-1}$) by the burial of hydrophobic surface. In contrast, the atomic solvation parameters derived by Eisenberg and McLachlan³⁶ predict that the contribution of the buried surface area of the Thr 157 γ -methyl group is only $0.04 \text{ kcal mol}^{-1}$. This could be correct if the indirect effects of the methyl group discussed above prove to be important. Further physical studies, including measurement of the hydrogen exchange rates¹³ of the main chain amide of

Asp 159, could help to distinguish between these disparate interpretations.

Comparison of the mutant proteins containing Ile and Val also leads to the conclusion that the δ -methyl group of Ile 157 destabilizes phage lysozyme by ~ 1.3 kcal per mole at 42 °C (Table 2). The δ -methyl group sequesters ~ 18 Å² of hydrophobic surface area in the folded protein¹⁰. The greater relative stability of the Val 157 variant suggests that this favourable hydrophobic contribution is offset by other effects. As the δ -methyl group of Ile 157 causes a shift of ~ 0.7 Å in the average position of the carboxylate of Asp 159, changes in electrostatic interactions could account for the observed destabilization.

This set of 13 amino-acid substitutions provides a view of the response of the lysozyme structure to chemically different amino acids at a single position on the protein surface (Fig. 2). Within the limitations of the resolution of the crystallographic data (1.7 Å), all thirteen new amino acids are accommodated by shifts in the protein and the solvent at or adjacent to residue 157. A similar result was seen for the substitution of Arg 96 by His on the surface of the protein⁹. In contrast, replacement of Gly 156 by Asp (ref. 37) and of the solvent-exposed Pro 86 by any of seven amino acids³⁸ results in structural shifts that propagate up to 20 Å across the protein surface.

Despite the strictly localized shifts associated with substitu-

tions of the Thr 157, the structures of the mutant proteins contain a number of unforeseen features, prominent among which is the new water molecule binding to the Gly 157 lysozyme (Fig. 1d). Additional bound solvent is also observed near Thr 155 or Asp 159 in response to several other substitutions. Arg 157 forms a new ion-pair. The mobility of the Ser 157 hydroxyl group is increased compared to Thr 157. The polypeptide backbone adjusts in response to the Asn 157 side chain (Fig. 1e). The δ -carbons of Leu and Ile can be superimposed on paper, but they are in very different positions in the X-ray crystal structures of the respective proteins (Fig. 2). Asp, Asn and His adopt different average conformations (Fig. 2), even though these are often considered conservative substitutions. These results emphasize the importance of high resolution structural studies for understanding the properties of mutant proteins.

We thank Drs D. C. Muchmore and F. W. Dahlquist for providing the lysozyme expression system; W. J. Becktel, W. A. Baase and J. A. Schellman for advice and facilities relating to the measurements of stabilities; L. H. Weaver, D. E. Tronrud, T. M. Gray, J. Bell and R. Faber for assistance with the crystallography and S. J. Remington for help with graphics. T.A. acknowledges the support of the Helen Hay Whitney Foundation. This work was supported in part by grants from the NIH (to B.W.M.) and NSF (to B.W.M. and J. A. Schellman).

Received 13 August; accepted 1 October 1987.

1. Privalov, P. L. *Adv. Protein Chem.* **33**, 167-241 (1979).
2. Schellman, J. A., Lindorfer, M., Hawkes, R. & Grütter, M. *Biopolymers* **20**, 1989-1999 (1981).
3. Privalov, P. L., Griko, Y. V., Venyaminov, S. & Kutysenko, V. P. *J. molec. Biol.* **190**, 487-498 (1986).
4. Elwell, M. & Schellman, J. *Biochim. biophys. Acta* **494**, 367-383 (1977).
5. Hawkes, R., Grütter, M. G. & Schellman, J. A. *J. molec. Biol.* **175**, 195-212 (1984).
6. Becktel, W. J. & Baase, W. A. *Biopolymers* **26**, 619-623 (1987).
7. Becktel, W. J. & Schellman, J. A. *Biopolymers* (in the press).
8. Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C. & Matthews, B. W. *Biochemistry* **26**, 3754-3758 (1987).
9. Grütter, M. G., Hawkes, R. B. & Matthews, B. W. *Nature* **277**, 667-669 (1979).
10. Grütter, M. G. *et al. J. molec. Biol.* **197**, 315-329 (1987).
11. Hudson, B. S. *et al. in Applications of Fluorescence in the Biomedical Sciences* (eds Taylor, D. L., Waggoner, A. S., Murphy, R. F., Laani, F. & Birge, R. R.) 159-202 (Liss, 1986).
12. Griffey, R. H., Redfield, A. G., Loomis, R. E. & Dahlquist, F. W. *Biochemistry* **24**, 817-822 (1985).
13. McIntosh, L. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 1244-1248 (1987).
14. Weaver, L. H. & Matthews, B. W. *J. molec. Biol.* **193**, 189-199 (1987).
15. Zoller, M. J. & Smith, M. *DNA* **3**, 479-488 (1984).
16. Alber, T. & Matthews, B. W. *Meth. Enzym.* **154**, 511-533 (1987).

17. Sanger, F., Nickelsen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
18. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
19. Wray, W., Boulukas, T., Wray, V. P. & Hancock, R. *Analyt. Biochem.* **118**, 197-203 (1981).
20. Perry, L. J. & Wetzel, R. *Science* **226**, 555-557 (1984).
21. Repaske, K. *Biochim. biophys. Acta* **22**, 189-191 (1956).
22. Remington, S. J. *et al. J. molec. Biol.* **118**, 81-98 (1978).
23. Schmid, M. G. *et al. Acta crystallogr.* **A37**, 701-710 (1981).
24. Rossmann, M. G. *J. appl. Crystallogr.* **12**, 225-239 (1979).
25. Jones, T. A. in *Crystallographic Computing* (ed. Sayre, D.) 303-317 (Oxford University Press, 1982).
26. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489-503 (1987).
27. Nemethy, G., Leach, S. J. & Scheraga, H. A. *J. Phys. Chem.* **70**, 998-1004 (1966).
28. Matthews, B. W., Nicholson, H. & Becktel, W. J. *Proc. natn. Acad. Sci. U.S.A.* **84** (in the press).
29. Janin, J., Wodak, S., Levitt, M. & Maigret, B. *J. molec. Biol.* **125**, 357-386 (1978).
30. Ponder, J. W. & Richards, F. M. *J. molec. Biol.* **193**, 775-792 (1987).
31. Klotz, I. M. & Franzen, J. S. *J. Am. chem. Soc.* **84**, 3461-3466 (1962).
32. Schellman, J. A. *Compt. Rend. Trav. Lab. Carlsberg Ser. Chim.* **29**, 223-229 (1955).
33. Susi, H., Timasheff, S. N. & Ard, J. A. *J. biol. Chem.* **239**, 3051-3054 (1964).
34. Creighton, T. E. *Biopolymers* **22**, 49-58 (1983).
35. Richards, F. M. A. *Rev. Biophys. Bioengng* **6**, 151-176 (1977).
36. Eisenberg, D. & McLachlan, A. D. *Nature* **319**, 199-203 (1986).
37. Gray, T. M. & Matthews, B. W. *J. biol. Chem.* **262** (in the press).
38. Alber, T. *et al. Science* (submitted).

LETTERS TO NATURE

Luminous arcs from galactic bow shocks

Mitchell C. Begelman* & Roger D. Blandford†

* Joint Institute for Laboratory Astrophysics, University of Colorado and National Bureau of Standards, Boulder, Colorado 80309-0440, USA

† Theoretical Astrophysics, 130-33 California Institute of Technology, Pasadena, California 91125, USA

Lynds and Petrosian¹ and Soucail *et al.*² have reported the discovery of circular arcs in the cores of two high-redshift ($z \sim 0.3$) clusters. The arcs are ~ 100 kpc in length, ~ 5 kpc in width, and have radii of curvature ~ 70 kpc ($H_0 = 75$ km s⁻¹ Mpc⁻¹). Giant elliptical galaxies are found near the centres of both arcs. The arcs have total luminosities of $\sim 10^{44}$ erg s⁻¹ and are unusually blue in colour. One explanation for the arcs is that they are regions of rapid star formation behind galaxy bow shocks, where our line of sight is tangent to the shock surface and approximately perpendicular to the galaxy velocity. Here we describe some of the

conditions under which bow shocks can form and point out some observationally testable implications of this model.

The inferred densities and temperatures of the X-ray emitting gas within several rich clusters imply radiative-cooling times short compared with the age of the Universe, but long compared with the local dynamical time³. Furthermore, X-ray observations of the centres of clusters⁴ indicate that there is a positive radial temperature gradient, a strong indicator of cooling. Third, several examples of irregular line-emitting filaments have been discovered near the centres of rich clusters⁵; these are generally supposed to be created by thermal instability in a cooling accretion flow⁶. From this, it has been concluded that the gas is settling into the cluster's gravitational potential well (which we parameterize using a core radius $r_c = 100 r_{c2}$ kpc and a central velocity dispersion $\sigma = 1,000 \sigma_3$ km s⁻¹). Under normal conditions, the inflow will be quite subsonic and the gas will have a temperature comparable with the virial temperature (in units of 10^6 K)

$$T_c \sim 70 \sigma_3^2 (r/r_c)^2 / \beta \quad (1)$$

where $\beta \sim 1$, the exact value depending upon the details of the model. Under these conditions, most galaxies will not form

strong bow shocks. But if conditions allow the gas to cool on a dynamic timescale, then the speed of sound in the cooling flow may become low enough to allow strong bow shocks to form.

The structure and dynamics of a cooling flow depends sensitively on entropy. For gas of density $\rho = 10^{-27} \rho_{-27} \text{ g cm}^{-3}$, molecular weight μ , temperature $T = 10^6 T_6 \text{ K}$ and entropy $S = (k/\mu m_p) \ln(T^{1.5}/\rho)$, (m_p is the proton mass) an approximate form for the cooling time for $10^5 \text{ K} \leq T \leq 4 \times 10^7 \text{ K}$ and solar abundance is⁷

$$t_{\text{cool}} \sim 80 T_6^{1.5} \rho_{-27}^{-1} \text{ Myr} \propto \exp(\mu m_p S/k) \quad (2)$$

essentially independent of temperature at a given entropy. If the gas has cooled to the point where it responds to the gravitational potential of the cluster, then the relevant dynamical time is essentially the free-fall time, given by

$$t_{\text{dyn}} \sim \frac{\pi r_c}{2\sqrt{3}\sigma} \sim 90 \left(\frac{r_{c2}}{\sigma_3} \right) \text{ Myr} \quad (3)$$

(independent of radius) for $r \leq r_c$. The ratio of these two timescales therefore depends exponentially on the entropy. The entropy, initially low if the gas is derived from galaxies, can subsequently either increase as a result of heating by shocks, stars and supernovae, or decrease because of radiative cooling. If, as we hypothesize, in the few clusters with arcs, the entropy is sufficiently low that $\eta = t_{\text{cool}}/t_{\text{dyn}} \sim 1$, the infall speed should become supersonic. But the motion is unlikely to be highly supersonic because irregularities in the flow inside the sonic radius will tend to amplify into shocks which will reheat the gas.

In general, cooling flows are believed to be thermally unstable⁸, and finite-amplitude negative entropy fluctuations should cool rapidly to form dense ($T \sim 10^4 \text{ K}$) clouds. These clouds, that become dynamically decoupled from the hot gas and move ballistically, are widely believed to produce low-mass stars⁹. If this process is efficient, it may lead to a crude form of self-regulation, with just enough material condensing into filaments at each radius to keep t_{cool} of the order of the inflow time¹⁰⁻¹². This self-regulation implies that

$$\rho_{-27} \sim 800 \eta^{-1} \beta^{-1.5} \sigma_3^4 r_{c2}^{-1} (r/r_c)^3 \quad (4)$$

The cool gas will presumably be mostly in pressure equilibrium at a density $\rho_{\text{cl}} \sim 6 \times 10^{-21} \sigma_3^6 r_{c2}^{-1} (r/r_c)^5 \text{ g cm}^{-3}$. The mass-accretion rate will be given by

$$\dot{M} \sim 2,000 T_6^{1.5} \sigma_3^2 r_2^3 \eta^{-1} r_{c2}^{-2} M_\odot \text{ yr}^{-1} \quad (5)$$

Note that almost all of the gas condenses out at large radius; very little is left to accrete onto the central galaxy.

Now suppose we have a cluster in which the gas can cool on the inflow time within the central $\sim 100 \text{ kpc}$. An elliptical galaxy moving supersonically ($v_{\text{gal}} \geq c_s = 1,300 \sigma_3 (r/r_c) \beta^{0.5} \text{ km s}^{-1}$) will accrete gas and be preceded by a blunt bow shock. This type of flow is amenable to numerical simulation, although existing calculations of similar accretion flows have not included the physical conditions that we envisage¹³⁻¹⁵. In particular, the galaxy is extended (with a soft potential), the entropy of the gas is probably variable and mass may condense out of the gas as it is compressed. The nature of the flow depends on the relationship of v_{gal} to the one-dimensional velocity dispersion, σ_{gal} , inside the galaxy. Simulations of the soft potential case with an adiabatic accreting gas¹⁶ suggest that a blunt bow shock forms when

$$c_s \leq v_{\text{gal}} \leq 2.5 \sigma_{\text{gal}}, \quad (6)$$

Typically, the stand-off distance of the bow shock is somewhat smaller than the radius of curvature, as is observed in both clusters. Essentially what happens is that the accreting gas is focused by the gravitational field of the galaxy, increasing the pressure enough to support a bow shock. (A galactic wind may also contribute to this increased pressure. But the mass-loss rate

is unlikely to be high enough for the wind to flow out to the observed arc radius¹⁷.) One consequence of the short dynamical and cooling timescale is that the radial inflow should re-establish itself very quickly after a galaxy has passed through the cluster core. We plan a numerical study of the non-adiabatic case with the intention of estimating the stand-off distance, shape and stability of bow shocks.

If we adopt these criteria, then a plausible choice of parameters is $\sigma_3 \sim 1$, $r_{c2} \sim 3$, $r_2 \sim 0.5$, and $\beta \sim \eta \sim 1$. In this case, $T_6 \sim 2$ and a galaxy with velocity dispersion σ_{gal} and speed in the range $200\text{--}700 \text{ km s}^{-1}$ will form a bow shock with Mach number $M \leq 3.5$. Giant elliptical galaxies in clusters are observed to have $\sigma_{\text{gal}} \leq 400 \text{ km s}^{-1}$, leaving ample room for condition (6) to be satisfied under the assumptions we have made. Restrictions on v_{gal} and conditions in the cooling flow are more stringent, however, and may explain the reported rarity of luminous arcs.

As the mixture of hot cluster gas and cool filaments crosses the bow shock, the pressure will increase by a factor $(5M^2 - 1)/4$, and will, we suggest, trigger impulsive gravitational collapse of the filaments to form massive stars, rather than the low mass stars postulated to be created in steady cooling flows. Our ignorance of the mechanism of star formation makes it pointless to try to justify this hypothesis. An initial mass function whose light is dominated by massive stars is necessary to account for both the colour of the shell and its surprising thinness. O stars ($\sim 20 M_\odot$) radiate blue light with an efficiency $\sim 10^{18} \text{ erg g}^{-1}$ and live for $\leq 10 \text{ Myr}$, so that with the envisaged post-shock speeds $\leq 200 \text{ km s}^{-1}$, they would indeed be confined to a shell of width $\leq 5 \text{ kpc}$. Thus, we propose that the arcs are rendered prominent just like spiral arms in galaxies. Making allowance for those portions of the shell which are presumably too faint to be seen, we find that the total star formation rate must be $\sim 1 M_\odot \text{ yr}^{-1}$, a few per cent of the total accretion rate. By contrast, the total gas kinetic energy flux crossing the shock is $\leq 4 \times 10^{43} \text{ erg s}^{-1}$, much too little to explain the observed luminosity as non-thermal synchrotron radiation or collisionally excited line emission.

We have shown that shock-induced massive star formation in a cooling flow can account for the luminous arcs if the gas is cool enough and if a giant galaxy moves fast enough through this gas. There are several direct observational tests and some indirect implications of this idea. First, the arc spectra must share redshifts with the clusters, although it will be very difficult to confirm this at the low light levels involved. Second, as emphasized by Braun and Dekel (preprint), the contrast between the mean arc surface brightness and that interior to the arcs must not be too large (typically ≤ 4) if the arcs are a thin shell seen in projection, whatever the shell's origin. We also expect there to be colour gradients across the arcs, with the bluer stars concentrated towards the outside. The massive stars should produce supernovae, mostly inside the arcs, at a rate of roughly one per ten years, far higher than the rate per unit mass associated with a normal stellar population. The spectra of H II regions should be easily detectable through a narrow band filter matched to a convenient, prominent emission line at the redshift of the cluster. We estimate that the region occupied by hot stars should contain at least several thousand million solar masses of cool gas at densities above several hundred cm^{-3} . This gas could easily intercept and reradiate much of the ultraviolet emission from the hottest stars, provided it has a sufficiently large covering factor. The line luminosity depends on the ionizing radiation flux, which in turn depends on the stellar mass function. By analogy with Minkowski's object, which is thought to be a starburst triggered by the impact of a jet¹⁸, we suggest that line emission may constitute as much as 10% of the optical continuum luminosity. Black hole and neutron star binaries may be present but are unlikely to dominate the X-ray emission ($\sim 10^{43} (\dot{M}/1 M_\odot \text{ yr}^{-1}) \text{ erg s}^{-1}$) from the diffuse cluster gas, although it is conceivable that the arc will be seen by AXAF (Advanced X-ray Astronomy Facility). The supernova blast waves, expanding into the surrounding interstellar medium, will

make the arc into a radio source. Adopting a radio radiative efficiency of $\sim 10^{12} \text{ erg g}^{-1}$, appropriate to local supernova remnants, we estimate that the total 5 GHz radio flux density of the arc should be $\sim 10 \mu\text{Jy}$, marginally detectable by the Very Large Array.

If our analysis of the conditions imposed on the cooling flow is correct, then the X-ray luminosity and spectrum of the cluster gas should be detectable and have the central concentration and inverted temperature profile characteristic of a cooling flow. In particular, by combining equations (2) and (3), we see that the gas will be able to cool on the dynamical timescale only if the core radius is fairly large and the X-ray luminosity substantial. As the creation of a significant bow shock depends upon the relative velocity between the inflowing gas and the galaxy, the arcs are more likely to be found on the opposite side of the galaxy from the cluster centre. This appears to be true of the two present examples. We also anticipate that the associated galaxies will have modest velocity differences (up to several hundred km s^{-1}) from the cluster mean.

In this model, we require that central giant elliptical and cD galaxies move with respect to the intracluster medium with speeds of typically half the cluster velocity dispersion. cD galaxies are generally thought to dominate the central potential wells of the clusters in which they are found, and indeed in some cases their radial velocities differ little from the cluster mean. However, some cDs possess wide-angle tail radio sources and appear to move with large relative velocities with respect to their surroundings¹⁹. These are sources in which apparently normal and well-collimated jets emerge along anti-parallel directions from the galaxy nucleus. At a distance $\sim 50 \text{ kpc}$, comparable with the radius of the observed arcs, the jets appear to decollimate and form two diffuse rails that bend towards the same side of the galaxy. One interpretation of wide-angle tails is that this apparent decollimation occurs when the jets leave a gaseous envelope co-moving with the galaxy and are exposed to the cluster gas blowing past it. A dense cooling flow can exert sufficient ram pressure to accomplish this in contrast to the case of an isothermal gas²⁰. It is therefore worthwhile to look for arcs upstream of wide-angle tails.

In conclusion, we find that it is possible to associate luminous galaxy arcs with massive star formation behind a galaxy bow shock and that several distinctive predictions can be made. It should be possible to confirm or refute this model with relatively straightforward observations.

We are grateful for the hospitality and stimulating environment of the Institute for Theoretical Physics, University of California, Santa Barbara, where this work was begun. We thank Avishai Dekel, John Hawley, Roger Romani, John Stocke and Dean Sumi for helpful discussions. M.C.B. thanks NSF, the A. P. Sloan Foundation, NASA, Rockwell International Corporation and Ball Aerospace Systems Division for support. RDB thanks NSF for support.

Received 18 May; accepted 24 August 1987.

- Lynds, R. & Petrosian, V. *Bull. Am. astr. Soc.* **18**, 1014 (1986).
- Soucaill, G., Fort, B., Mellier, Y. & Picat, J. P. *Astr. Astrophys.* **172**, L14-L16 (1987).
- Forman, W. & Jones, C. A. *Rev. Astr. Astrophys.* **20**, 547-585 (1982).
- Mushotzky, R. F., Holt, S. S., Smith, B. W., Boldt, E. A. & Serlemitsos, P. J. *Astrophys. J.* **244**, L47-L51 (1981).
- Cowie, L., Hu, E., Jenkins, E. & York, D. *Astrophys. J.* **272**, 29-47 (1983).
- Sarazin, C. *Rev. Mod. Phys.* **58**, 1-115 (1986).
- McKee, C. F. & Cowie, L. L. *Astrophys. J.* **215**, 213-225 (1977).
- Mathews, W. G. & Bregman, J. N. *Astrophys. J.* **224**, 308-319 (1978).
- Fabian, A. C., Nulsen, P. E. & Canizares, C. R. *Nature* **310**, 733-740 (1984).
- Nulsen, P. E. J., Stewart, G. C. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **208**, 185-195 (1984).
- Begelman, M. C. *Nature* **322**, 614-615 (1986).
- White, R. & Sarazin, C. L. *Astrophys. J. Suppl. Ser.* (in the press).
- Hunt, R. *Mon. Not. R. astr. Soc.* **154**, 141-165 (1971).
- Hunt, R. *Mon. Not. R. astr. Soc.* **188**, 83-91 (1979).
- Fryxell, B. A., Taam, R. E. & McMillan, S. L. W. *Astrophys. J.* **315**, 536-554 (1987).
- Gisler, G. R. *Astr. Astrophys.* **51**, 137-150 (1976).
- Takeda, H., Nulsen, P. E. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **208**, 261-278 (1984).
- Van Brughel, W., Filippenko, A., Heckman, T. & Miley, G. *Astrophys. J.* **293**, 83-93 (1985).
- Burns, J. O. *Mon. Not. R. astr. Soc.* **195**, 523-533 (1981).
- Eilek, J. A., Burns, J. O., O'Dea, C. P. & Owen, F. N. *Astrophys. J.* **278**, 37-50 (1984).

A magnetic 'glassy' state in $\text{La}_2\text{CuO}_{4-y}$

Madoka Tokumoto, Yoshikazu Nishihara, Kunihiko Oka & Hiromi Unoki

Electrotechnical Laboratory, Tsukuba Research Center, Ibaraki 305, Japan

Since the discovery of high- T_c superconductivity in $\text{La}_{2-x}\text{Ba}_x\text{CuO}_{4-y}$ and related perovskite systems¹⁻⁵, the magnetic properties of the parent compound $\text{La}_2\text{CuO}_{4-y}$ have been of great interest, particularly with respect to the possibility that the mechanism for superconductivity in these oxide superconductors may involve antiferromagnetism⁶⁻⁸. In undoped $\text{La}_2\text{CuO}_{4-y}$, the magnetic susceptibility shows a cusp and antiferromagnetic ordering has been confirmed by neutron scattering⁹, nuclear magnetic resonance^{10,11} and Mössbauer spectroscopy¹². Yet the magnetic properties have not been fully reconciled with the proposed spin structure^{10,13}. Here we report static magnetic susceptibility measurements on a single crystal of undoped $\text{La}_2\text{CuO}_{4-y}$ that reveal the existence of a new magnetic state. In low magnetic field, we observed a magnetic 'glassy' behaviour below the Néel temperature T_N ($\sim 270 \text{ K}$), indicating that the ground state of this material is not a simple antiferromagnetic state. These observations may have some relevance to the origin of the high- T_c superconductivity.

Although most of the measurements reported so far have been made on polycrystalline ceramics, the preparation and characterization of high-quality single crystals are now essential for the proper understanding of this quasi-two-dimensional system, including its anisotropy. Single crystals as large as $15 \times 20 \times 3 \text{ mm}$ of undoped $\text{La}_2\text{CuO}_{4-y}$ were grown by the top-seeded solution growth method described in ref. 14. The static magnetic susceptibility measurements were made with a commercial SQUID susceptometer (SHE Model 905). An as-grown single crystal ($5 \times 5 \times 1 \text{ mm}$, 0.28 g) was mounted with the c axis either parallel or perpendicular to the field direction. (For simplicity, we use tetragonal crystallographic nomenclature, so that the c axis is perpendicular to the CuO_2 plane.)

Figure 1a shows the temperature dependence of the static magnetic susceptibility $\chi(T)$ measured with the magnetic field of 10 kOe parallel and perpendicular to the c axis. When the field is parallel to the c axis, the peak in the susceptibility corresponding to T_N occurs at 270 K ; when the field is perpendicular, a smaller peak is observed at the same temperature. A completely different behaviour was observed when the magnetic susceptibility measurement was made with much smaller magnetic fields. Figure 1b shows the temperature dependence of the static magnetic susceptibility of the same sample with an applied field of 100 Oe . When the temperature is lowered with the magnetic field parallel to the c axis, new features appear in the susceptibility. First, a prominent peak grows up at $\sim 240 \text{ K}$ —that is, at a somewhat lower temperature than the peak in Fig. 1a—and secondly, a sharp decrease in the susceptibility is observed below 115 K . Comparison with the data at 10 kOe (Fig. 1a) indicates that the difference comes from an enormous enhancement of the susceptibility below T_N . The susceptibility measured in a field (100 Oe) perpendicular to the c axis shows almost the same temperature dependence as the data at 10 kOe , but the magnitude is somewhat larger. Note that the peak in $\chi(T)$ indicating a new magnetic phase is observed only with the magnetic field parallel to the c axis, whereas the peak corresponding to T_N is seen for both parallel and perpendicular configurations (Fig. 1a). Suppression of the peak in $\chi(T)$ by application of higher magnetic fields suggests the existence of a spin-glass-like magnetic state.

The zero-field-cooled (ZFC) susceptibility χ^{ZFC} was measured by heating in a magnetic field after zero-field cooling to 6 K ; the field-cooled (FC) susceptibility χ^{FC} was measured by apply-

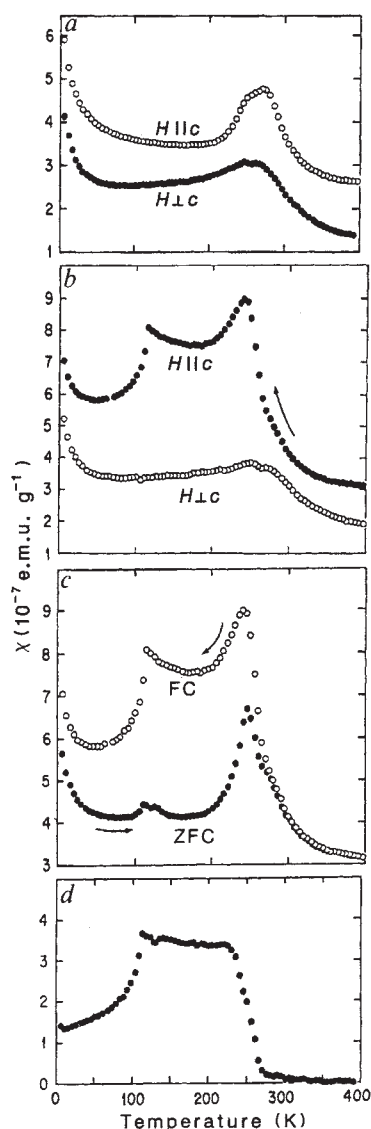


Fig. 1 Temperature dependence of magnetic susceptibility in $\text{La}_2\text{CuO}_{4-y}$. *a*, High-field susceptibility; applied field, 10 kOe. *b*, Low-field susceptibility; applied field, 100 Oe. *c*, Field-cooled (χ^{FC}) and zero-field-cooled (χ^{ZFC}) susceptibility; applied field, 100 Oe, parallel to *c* axis. *d*, $\chi^{\text{FC}} - \chi^{\text{ZFC}}$; applied field as in *c*.

ing a measuring field at high temperatures and then cooling in the field. Figure 1c shows the ZFC and FC susceptibilities measured with the field parallel to the *c* axis, and Fig. 1d shows the temperature dependence of the glassy component of magnetic susceptibility ($\chi^{\text{FC}} - \chi^{\text{ZFC}}$). The glassy component begins to increase just below T_N (~270 K), saturates at ~230 K and shows a plateau at temperatures between 230 and 115 K. Note the very small temperature dependence in this temperature region. Below 115 K it shows a sharp decrease, indicating the appearance of another new phase for which the excess glassy component of magnetic susceptibility along the *c* axis decreases with decreasing temperature, although it seems to remain finite even at $T=0$. A cusp of susceptibility at ~110 K has also been observed in a polycrystalline sample¹².

The characteristic aspect of this magnetic glassy state is the presence of two successive transitions associated with a single component of spin susceptibility. The decrease in the glassy component below 115 K suggests the presence of freedom in spin orientation along the *c* axis at temperatures well below T_N or the onset of glassy behaviour. This cannot be explained in the framework of an ordinary spin glass; thus, the spins are not

frozen but are probably fluctuating in this 'magnetic glassy state'. A possible origin for the superconductivity could be, for example, the propagation of *p*-holes of oxygen, which break copper-copper antiferromagnetic bonds^{7,8}. But a more quantitative theory, consistent with the spin structure in the antiferromagnetic phase⁹, is needed to explain the observed magnetic properties of $\text{La}_2\text{CuO}_{4-y}$.

We acknowledge encouragement and support by T. Ishiguro, S. Maekawa, Y. Kimura and K. Murata, and discussions on spin glasses with Y. Miyako.

Note added in proof: We have recently learned of neutron diffraction observations¹⁵ of the magnetic structure of La_2CuO_4 , which do not yield evidence of any transition at 115 K. It may be that the neutron measurement is not sensitive enough to detect the small change in susceptibility.

Received 9 October; accepted 19 October, 1987.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys.* **B64**, 189-193 (1986).
2. Uchida, S., Takagi, H., Kitazawa, K. & Tanaka, S. *Jap. J. appl. Phys.* **26**, L1-L2 (1987).
3. Wu, M. K. *et al.* *Phys. Rev. Lett.* **58**, 908-910 (1987).
4. Zhao, Z. *et al.* *Kexue Tongbao* **32**, 412-414 (1987).
5. Hikami, S., Hirai, T. & Kagoshima, S. *Jap. J. appl. Phys.* **26**, L314-L315 (1987).
6. Anderson, P. W. *Science* **235**, 1196-1198 (1987).
7. Emery, V. J. *Phys. Rev. Lett.* **58**, 2794-2797 (1987).
8. Hirsch, J. E. *Phys. Rev. Lett.* **59**, 228-231 (1987).
9. Vaknin, D. *et al.* *Phys. Rev. Lett.* **58**, 2802-2805 (1987).
10. Kitaoka, Y., Hiramatsu, S., Ishida, K., Kohara, T. & Asayama, K. *J. phys. Soc. Japan* **56**, 3024-3027 (1987).
11. Watanabe, I. *et al.* *J. Phys. Soc. Japan* **56**, 3028-3031 (1987).
12. Nishihara, Y., Tokumoto, M., Murata, K. & Unoki, H. *Jap. J. appl. Phys.* **26**, L1416-L1418 (1987).
13. Fukuda, K., Sato, M., Shamoto, S., Onoda, M. & Hosoya, S. *Solid St. Commun.* **63**, 811-813 (1987).
14. Oka, K. & Unoki, H. *Jap. J. appl. Phys.* **26**, L1590-L1592 (1987).
15. Shirane, G. *et al.* *Phys. Rev. Lett.* **59**, 1613-1616 (1987).

Possible role of Cu^{2+} - Cu^{4+} pairs in the superconductivity of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ from electron spin resonance observations

M. D. Sastri, A. G. I. Dalvi, Y. Babu, R. M. Kadam, J. V. Yakhmi & R. M. Iyer

Chemical Group, Bhabha Atomic Research Centre, Trombay, Bombay 400 085, India

The recent discovery¹⁻³ of high-temperature superconductivity in the $\text{Y}(\text{La})\text{-Ba-Cu-O}$ system has generated intense experimental and theoretical activity. On the theoretical side, it has revived the discussion of different modes of pairing in solids^{4,5}. Here we report some important features of the low-field electron spin resonance (ESR) spectrum in the superconducting phase of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, giving evidence for the dimerization of electrons on copper pairs. An intense low-field resonance appears below the critical temperature, T_c , and exhibits unresolved hyperfine structure at temperatures below 70 K. The resonance field was found to be dependent on both temperature and microwave frequency. These are well-established features characteristic of copper pairs formed due to an exchange interaction. It appears that the individual spins responsible for superconductivity are formed and dimerized only at T_c , with their intensity sharply increasing below T_c . This observation appears to be a pointer towards the possibility of the disproportionation $2\text{Cu}^{3+} \rightarrow \text{Cu}^{2+} + \text{Cu}^{4+}$ in adjacent octahedra, facilitated by coupling to local e_g vibrations in the superconducting phase.

Superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ was prepared and characterized by the methods described in ref. 3; X-ray diffraction confirmed the presence of a single phase. The ESR measurements were made on a pellet measuring $1 \times 2 \times 4$ mm, on an x-band spectrometer (Varian V-4502), and the temperature was varied between 300 and 35 K using a CTI-Cryosystem closed-

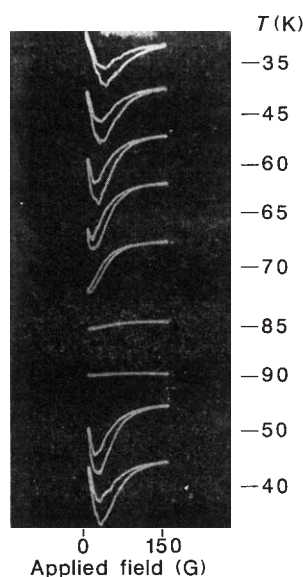


Fig. 1 Oscilloscope traces of the ESR spectrum of a sample ('A') exhibiting zero resistance at 103 K, cooled in zero magnetic field at different temperatures. Partially resolved hyperfine structure can be seen at low temperatures, as well as a shift in the line position with temperature.

cycle helium refrigerator. The temperature was varied using a Palm Beach Cryophysics model 4025 temperature controller with built-in silicon diodes for temperature measurements.

The ESR spectrum observed in the normal state was characteristic of Cu^{2+} ($g_{\parallel} = 2.176 \pm 0.001$, $B_{\parallel} \approx 65$ G; $g_{\perp} = 2.055 \pm 0.001$, $B_{\perp} \approx 0$ G) with an unpaired electron in the $d_{x^2-y^2}$ orbital. The intensity of the spectrum is relatively weak, and is similar to that of a dilutely doped Cu^{2+} sample with 10^{16} spins per cm^3 . This is rather surprising in view of the large concentration of Cu^{2+} expected in this compound⁴. As the sample was cooled through the normal-to-superconducting phase transition, the ESR spectrum observed at room temperature was nearly unaffected except for a slight reduction in intensity. On the other hand, at temperatures below 100 K a new resonance was observed at zero field and its intensity increased with decreasing temperature. This resonance was recently reported by Bhatt *et al.*⁷ at 77 K. The intensity of the zero-field line at 77 K (at $g \approx 2.0$ – 2.2) is at least three orders of magnitude higher than the resonance observed at 100 K.

Intense ESR absorption is detected at low fields (0–200 G), and exhibits magnetic hysteresis effects. Its resonance frequency (field) and intensity are found to be temperature-dependent (see Figs 1 and 2, respectively). Figure 1 shows the oscilloscope traces recorded at temperatures between 35 and 100 K on a sample cooled at zero field, to avoid any possible flux trapping in the superconducting phase. (Oscilloscope traces were found to be less susceptible than chart recordings to the noise that arises from the high sensitivity of intensity to small temperature fluctuations.) Very clear, partially resolved hyperfine structure can be seen in Fig. 1, with increasing resolution at lower temperatures, and a temperature-dependent shift of the resonance field towards zero field as T_c was approached from below.

In addition to the above, a sharp field-independent absorption was observed. This was found to be sharpest at low (100 kHz) modulation amplitude, and broadened as the modulation amplitude increased. Finally, when the sample was heated at 873 K in a nitrogen atmosphere for 12 h (Fig. 2b), the zero-field line appeared at $T \leq 60$ K, with a greatly reduced intensity. This is consistent with the expected⁸ decrease of T_c due to oxygen depletion.

These observations clearly suggest that the low-field resonance observed below 100 K arises only due to the electrons respon-

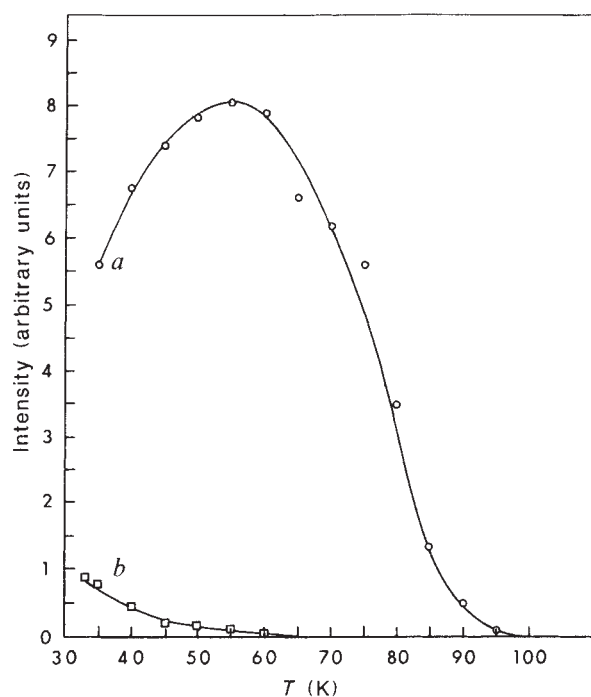


Fig. 2 Temperature dependence of the intensity of the zero-field line in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$: a, for sample A; b, for sample A after heating in nitrogen at 873 K for 12 h.

sible for the superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, and that the field-independent microwave absorption is also characteristic of the superconducting phase. In recent ESR studies of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, Durny *et al.*⁹ and Blazey *et al.*¹⁰ reported microwave absorption near zero field and attributed it to the a.c. diamagnetic susceptibility characteristic of clusters of superconducting grains in oxide superconductors. Some of the important features of the low-field resonance observed in our work—specifically, the temperature dependence of the resonance field and the existence of partially resolved structure, in samples cooled in zero magnetic field (Fig. 1), and the peaking of the intensity versus temperature curve between 50 and 60 K—do not seem to fit into this model. On the other hand, these features are strikingly similar to those reported for copper dimers in copper acetate¹¹ and copper propionate¹², where intense zero-field resonance was reported. The zero-field resonance arises from the formation of antiferromagnetically coupled dimers. The ESR transition occurs from the $M_s = 0$ to $M_s \pm 1$ states, separated by 0.3 cm^{-1} , of the low-lying excited triplet state, sufficiently populated for the observation of ESR absorption in the 30–80-K range. The partially resolved structure, presumably due to ^{63}Cu – ^{63}Cu and ^{65}Cu – ^{65}Cu hyperfine interaction, and the observed maximum in intensity at ~ 55 K (symptomatic of thermal depopulation of the excited state at low temperature) lend support to this mechanism. Furthermore, the clear temperature dependence of the line position (Fig. 1) can be satisfactorily accounted for in this model, as the exchange coupling responsible for dimer formation is temperature-dependent and is reflected in the temperature dependence of resonance frequency. As shown in ref. 11, the near-zero-field absorption for copper dimers is markedly dependent on the microwave frequency used. Experimental limitations prevented us from studying the ESR of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ at K/Q-band frequencies, but our experiments at 9.405 GHz using a Varian rectangular multi-purpose cavity (Fig. 3a), and at 9.030 GHz using a Varian cylindrical cavity (Fig. 3b), both at 77 K (sample cooled in zero field), showed a clear shift in the line position towards lower fields as the microwave frequency was reduced.

The nature of the copper dimers observed here is distinctly

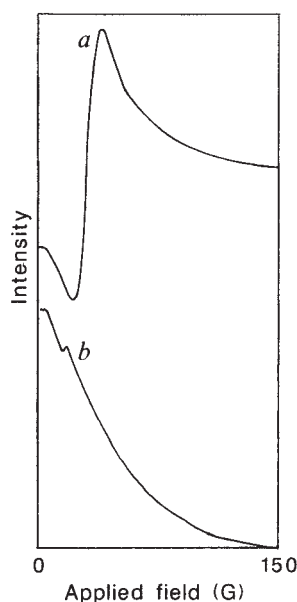


Fig. 3 ESR spectrum of sample A, cooled to 77 K in zero field in two different cavities: *a*, rectangular cavity resonating at 9.405 GHz; *b*, cylindrical cavity at 9.03 GHz.

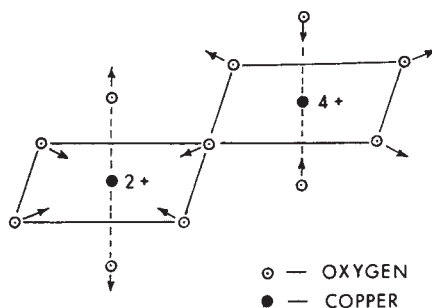


Fig. 4 e_g Vibrations of octahedral units. The elongation and compression of adjacent octahedra facilitate the disproportionation of Cu^{3+} . $\odot$, Oxygen; $\bullet$, copper.

different from that in copper acetate and propionate, because the individual spin states taking part in dimer formation do not exist in the normal phase of the compound. Their formation and concomitant pairing (on a timescale longer than the 10^{-10} -s timescale of ESR) appears to characterize the superconducting phase. If they had existed either as isolated spins or paired dimers in the normal phase, their ESR should have been detected with much greater intensity in the normal phase.

The high intensity of pair spectrum observed below T_c therefore suggests that pair formation has to be preceded by some charge disproportionation and simultaneous pairing. This can be either of the following: $2\text{Cu}^+ \rightarrow \text{Cu}^{2+} + \text{Cu}^0$ eventually forming Cu^{2+} - Cu^{2+} pairs and Cu^0 clusters, or, $2\text{Cu}^{3+} \rightarrow \text{Cu}^{4+} + \text{Cu}^{2+}$ forming Cu^{4+} - Cu^{2+} pairs. So far no evidence for Cu^0/Cu -clusters has been observed by us or by others. An interesting example of the latter possibility, in the present context, is Cu^{2+} and Cu^{4+} in adjacent sites, with Cu^{2+} in an elongated-octahedral (or square-planar) configuration and Cu^{4+} (low-spin $3d^7$) in a compressed octahedron, as the unpaired electron occupies the $x^2 - y^2$ orbital at both sites.

The present findings of the creation and pairing of spins below T_c point towards this possibility; furthermore, as shown in Fig. 4, local ' e_g ' vibrations indeed produce a coordination around two adjacent octahedra sharing one oxygen. Charge disproportionation of the type $2\text{Cu}^{3+} \rightarrow \text{Cu}^{4+} + \text{Cu}^{2+}$, with an electron transfer from the Cu^{3+} at the compressed octahedral site to that

at the elongated one creates $d_{|x^2-y^2|}^1 (\text{Cu}^{4+})$ and $d_{|x^2-y^2|}^1 (\text{Cu}^{2+})$. This facilitates the stabilization of dimers, as this configuration would have lower energy than that of two Cu^{3+} ions in adjacent octahedra. This results in valence fluctuations of copper between +4 and +2 on the timescale of the e_g vibration. When this timescale exceeds 10^{-10} s, the valence fluctuations can be observed as isolated pairs in the ESR spectrum. (Yet it is not possible to identify isolated Cu^{4+} or Cu^{2+} , as they exist only as pairs below T_c .) As pointed out in ref. 13, the singlet to triplet transition becomes allowed in the case of dimers made up of dissimilar ions. This can account for the additional peaks reported¹⁴ near 255 cm^{-1} in the optical reflection experiments. We therefore believe that the vibration-assisted disproportionation $2\text{Cu}^{3+} \rightarrow \text{Cu}^{4+} + \text{Cu}^{2+}$ (in adjacent octahedra) and formation of $(\text{Cu}^{4+}, 3d_{x^2-y^2} \uparrow) - (\text{Cu}^{2+}, 3d_{x^2-y^2} \downarrow)$ through a bridging oxygen are associated with the onset of superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$.

Received 20 August; accepted 8 October 1987.

1. Bednorz, J. G. & Müller, K. A. Z. *Phys.* B64, 189 (1986).
2. Wu, M. K. *et al.* *Phys. Rev. Lett.* 58, 908-911 (1987).
3. Gopalakrishnan, J. K., Yakhmi, J. V. & Iyer, R. M. *Nature* 327, 604-605 (1987).
4. Anderson, P. W. *Science* 235, 1196-1198 (1987).
5. Emery, V. J. *Phys. Rev. Lett.* 58, 2794-2797 (1987).
6. David, W. I. F. *et al.* *Nature* 327, 310-312 (1987).
7. Bhatt, S. V., Ganguly, P. & Rao, C. N. R. *Pramana-J. Phys.* 28, L425-427 (1987).
8. Johnston, D. C. *et al.* Pap. presented at Am. chem. Soc. Meet. New Orleans (31 August-4 September, 1987).
9. Durny, R. *et al.* *Phys. Rev. B* 36, 2361-2363 (1987).
10. Blazey, K. W. *et al.* *Phys. Rev.* (in the press).
11. Bleaney, B. & Bowers, K. D. *Proc. R. Soc. A* 214, 451-465 (1952).
12. Abe, H. *J. phys. Soc. Japan* 13, 987-997 (1958).
13. Abragam, A. & Bleaney, B. *Electron Paramagnetic Resonance of Transition Ions* (Clarendon, Oxford, 1970).
14. Wrobel, J. M., Wang, S., Gygas, S., Clayman, B. P. & Peterson, L. K. *Phys. Rev. B* 36, 2368-2370 (1987).

High-strength ceramics through colloidal control to remove defects

N. McN. Alford, J. D. Birchall & K. Kendall

ICI plc, PO Box 8, Runcorn, Cheshire WA7 4QD, UK

Although the theoretical tensile strength of alumina is $\sim 46\text{ GPa}$ (ref. 1), conventional powder processing rarely gives products better than 0.5 GPa because large defects, ~ 50 - $100\text{ }\mu\text{m}$ in size, are invariably trapped during powder compaction^{2,3}. Thin-film, fibrous, melt-processed and vapour-deposited ceramics can achieve higher strengths of several GPa (refs 4-6), but such processes are much more expensive than simple sintering of powder compacts. Here we show how colloidal control of powders to remove agglomerates can be used to give high-strength sintered ceramics: for example, a commercial alumina powder can be improved from 0.37 to 1.04 GPa in bend strength.

It is well known that the overwhelming problems of ceramic materials are large defects and low toughness⁷. Together, these factors give low tensile strength, σ , for a stretched ceramic sheet of toughness K_{Ic} containing an edge crack of length c extending through the thickness of the sheet, according to the fracture mechanics equation^{8,9}

$$\sigma = K_{Ic} / 1.12\sqrt{\pi c} \quad (1)$$

Although the toughness, K_{Ic} in this equation, has been improved by incorporating fibres^{10,11} or adding transformable zirconia^{12,13}, attention has also been drawn to the reduction of flaw size, c , to achieve higher strength. In particular, Lange and co-workers¹⁴⁻¹⁷ have studied colloidal processes for removing flaws. But there has yet been no explicit verification of fracture mechanics in terms of process-related defects, and no one has

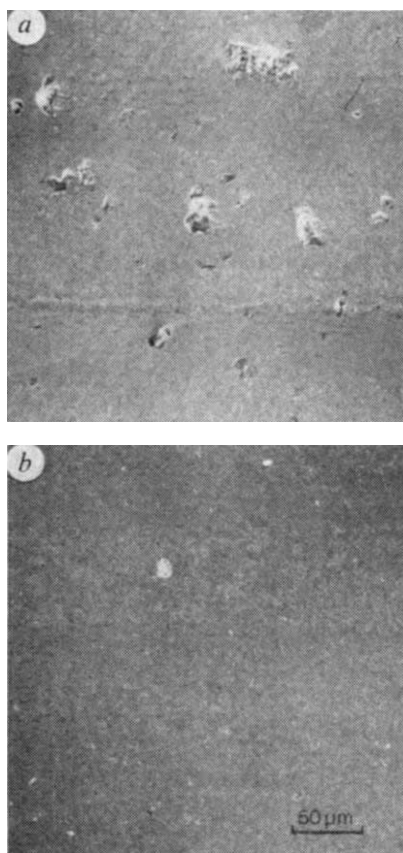


Fig. 1 *a*, Polished surface of sintered alumina made by powder pressing, illustrating the large defects in the material. *b*, Polished surface of sintered alumina made by mixing in polymer solution, followed by intense shearing. The defects were much reduced in size by this process.

succeeded in reducing flaws $\leq 20 \mu\text{m}$ in monolithic ceramic bodies, even by hot pressing methods. We demonstrate a direct correlation between measured process flaws and strength, associating such flaws with powder agglomerates. Then it is shown that the shearing of these agglomerates down to $10 \mu\text{m}$ and even down to $5 \mu\text{m}$ by novel colloidal processing gives remarkable strength enhancement in alumina and titania.

A key problem is the measurement of defects. Two methods were developed. The first relied on polishing the ceramic, photographing many areas of the surface in light reflected normally, and measuring the maximum dimension of black flaw-like regions with an image analyser. Figure 1*a* shows a typical area of a conventional alumina sample, with flawed patches. Plotting the largest flaws out of several thousand counts on a histogram (Fig. 2, open bars) shows the many defects ($0.2 \text{ v\%}$) between 40 and $120 \mu\text{m}$ in a conventional alumina (Reynolds HPDBM grade), plastic mixed, pressed into sheet, burned out and sintered at $1,550^\circ\text{C}$ for 1 h, density 3.90, grain size $6 \mu\text{m}$). To prove that the original powder contained agglomerates of this size, an ultrasonically dispersed mix of the alumina in water was filtered through a $40 \mu\text{m}$ mesh: 0.1% of the powder would not pass through the mesh.

A second method of measurement was necessary because microscopy cannot resolve very fine cracks, which may act as invisible flaws. This method used grain size as a marker for flaw size. The maximum grain size was measured optically. The grains were then grown by heat-treating the sample and the new maximum grain size determined. This procedure was repeated until the strength began to fall with increasing grain size (Fig. 4). The maximum grain size at this juncture was taken as the process-related defect in the sample^{18,19}. This grain-size com-

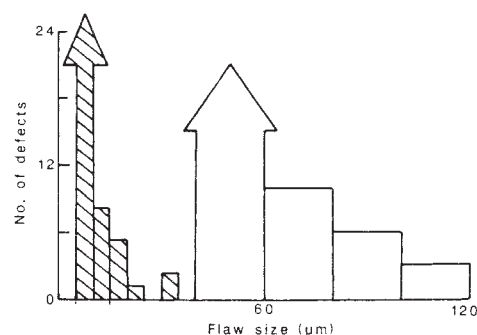


Fig. 2 Histograms of defects measured by microscopy of polished alumina surfaces. Open histogram, alumina made by conventional processing; shaded histogram, alumina made by dispersing in polymer solution followed by intense shearing (several thousand flaws were counted but only the largest flaws are shown for clarity).

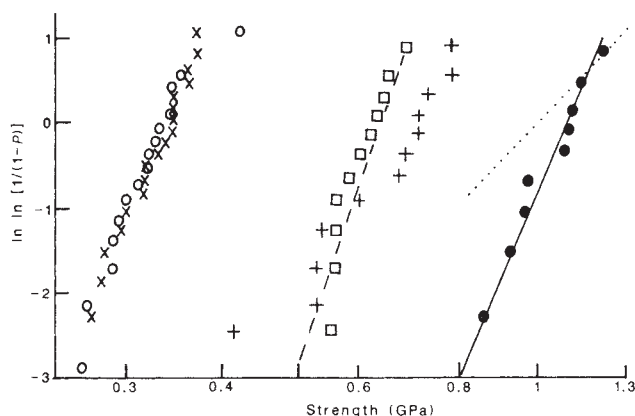


Fig. 3 Weibull plots of scatter in alumina strengths measured by bending, and predicted from the flaw size measurements: $\circ$, bending tests on Reynolds alumina processed conventionally; $\times$, calculated from flaw sizes by equation (2) for conventional Reynolds alumina; $\square$, bending tests on Reynolds alumina sheared in polymer solution, 2 mm diameter extrudate; $+$, calculated from flaw sizes measured microscopically; $\bullet$, bending tests on AKP30 alumina powder after intense shearing in polymer solution, 1 mm diameter extrudate; $\cdots$, calculated from microscopy.

parator method gave a defect size of $100 \mu\text{m}$ for the Reynolds HPDBM alumina. For comparison, the microscopy gave a mean flaw size of $83 \mu\text{m}$ for the 20 largest defects found in 20 areas 1 mm square on polished sections. In addition, when these defects were used to estimate the strength of the alumina using the equation (where K_{1c} is $4.0 \text{ MPa m}^{1/2}$ for alumina)

$$\sigma = 0.737 K_{1c} / \sqrt{c} \quad (2)$$

similar to that applying to edge flaws which do not extend through the sample thickness²⁰, excellent agreement was obtained with three-point bend strength measurements of these samples plotted on a Weibull graph²¹. This correlation is shown in Fig. 3, left curve.

Having measured the process-related defects, related them to powder agglomerates and correlated flaw length with bending strength, we then studied the control of strength through colloidal means. In the first experiments, a commercial alumina (Reynolds HPDBM) was mixed with a 40 wt % aqueous solution of polyvinylalcohol-acetate (Nippon Gohsei KH17s) using a two-roll mill. This dough-like material was pressed into 2-mm-thick flat sheets, which were dried at 80°C , cut into strips 5 mm wide and heated at 1°C min^{-1} to remove the polymer before firing at $1,550^\circ\text{C}$ for 1 h to sinter them. The density was 98% of theoretical and the maximum grain size was $6 \mu\text{m}$. Three-point

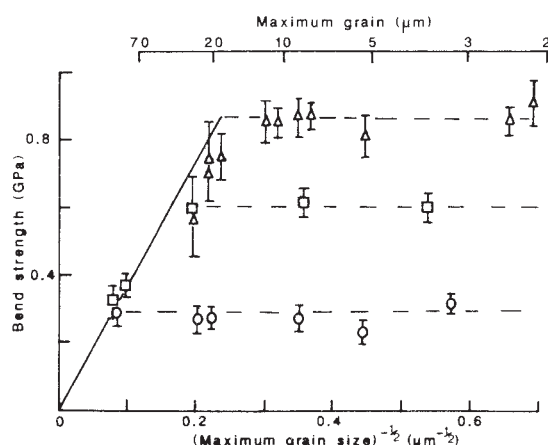


Fig. 4 Grain-size comparison results for alumina: $\circ$, conventionally processed alumina; $\square$, alumina sheared in polymer solution, 2 mm diameter extrudate; $\triangle$, AKP30 alumina sheared in polymer solution, 1 mm diameter extrudate. The maximum grain size when strength became dependent on grain size was taken to be the defect size in the material.

bend testing on a 16-mm span showed that the strength was 300–400 MPa immediately after firing, as expected from the 100- μm defects found microscopically in the material. Such strength and defects were also found for samples of this powder compacted by pressing and by slip casting. The conclusion was that agglomerates were persisting from the powder, through the compaction and sintering steps, to cause flaws in the final product.

To test this argument, the same plastic mix of Reynolds alumina was ram-extruded through an orifice 2 mm in diameter and 2 mm long, at a nominal wall shear stress of 10 MPa, and the extruded rods were sintered and tested as before, giving material of 99% of theoretical density with a grain size of 5 μm . The plastic extrudate was seen to be less lumpy than before, suggesting that agglomerates were being disrupted by the high shear stress. This improved structure was verified on polished sections (Fig. 1b). In the histogram of defect size, the curve was shifted to the left (Fig. 2, shaded curve), showing a reduced volume fraction of flaws, 5×10^{-6} , and predicting increased strength from equation (2) (Fig. 3, +). Three-point bend tests, ex-furnace, gave an encouraging fit to these predictions (Fig. 3, $\square$). Grain growth measurements supported these observations, giving a process flaw size of 37 μm compared with 21 μm by microscopy. The mean bend strength was 601 MPa, almost double that obtained without shearing the mix.

Extending this procedure, the extrusion orifice was further reduced to 1 mm diameter, raising the wall shear stress to 20 MPa. Microscopy showed a slight reduction in defect size, the strength rising only to 760 MPa, which was not increased by extra shear, indicating that the Reynolds powder contained a significant number of very hard lumps. Such hard lumps can be explained by studies that have shown that perfect, well-packed alumina agglomerates can exhibit strengths of several hundred MPa owing to surface forces between grains^{22,23}. Similar results were obtained with Alcoa CT3000SG alumina powder.

Selecting a powder with weak agglomerates improved the results considerably. Sumitomo alumina (AKP30), when processed in the same way, gave rods 1 mm in diameter with a mean strength of 1,042 MPa as indicated by the right-hand points in Fig. 3. Microscopy of polished samples showed a reduced flaw size, and when equation (2) was used to predict strength the dotted line in Fig. 3 was obtained. This gave reasonable agreement with the strength measurements, but of lower Weibull modulus. The mean large-flaw size was 15 μm compared with 12 μm by grain growth. When simply mixed on a two-roll mill

the strength was only 366 MPa. Similarly, pigment-grade titania (RCR2, Tioxide) gave defects 40 μm in length when compacted by two-roll mixing, but these were reduced to 3.9 μm by extrusion through an orifice 1 mm in diameter. These results show that it should be possible to produce alumina from powder to give a strength ~ 1.5 GPa, comparable with present-day reinforcing fibre, and substantially higher than previous reports for monolithic alumina.

We have shown that the defects causing weakness in ceramics made from commercial alumina powders are visible on polished sections and are associated with agglomerates in the powder. These defects may be removed by dispersing colloiddally in polymer solution, followed by shearing at high stress to break down the lumps. Remarkably low flaw sizes have been attained by this novel route^{24,25}.

Received 24 April; accepted 20 August 1987.

1. Kelly, A. & Macmillan, N. H. *Strong Solids*, ch. 1 (Clarendon, Oxford, 1986).
2. Alford, N. McN., Birchall, J. D. & Kendall, K. *Mater. Sci. Technol.* **2**, 329–336 (1986).
3. Lange, F. F. *J. Am. ceram. Soc.* **66**, 396–398 (1983).
4. Mallinder, F. P. & Proctor, B. A. *Phil. Mag.* **13**, 197–208 (1966).
5. Soltis, P. J. *Bull. Am. phys. Soc.* **10**, 163 (1965).
6. Birchall, J. D., Bradbury, J. A. A. & Dinwoodie, J. in *Handbook of Composites* Vol. 1 (eds Watt, W. & Perov, B. V.) 115–154 (Elsevier, Amsterdam, 1985).
7. Abe, H. *Bull. Am. ceram. Soc.* **64**, 1594–1596 (1985).
8. Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163–198 (1920).
9. Koiter, W. T. *Proc. R. Neth. Acad. Sci. B* **59**, 354–374 (1956).
10. Brennan, J. J. & Prew, K. M. *J. Mater. Sci.* **17**, 2371–2383 (1982).
11. Sambell, R. A. J., Bowen, D. H. & Phillips, D. C. *J. Mater. Sci.* **7**, 663–675 (1972).
12. Garvie, R. S., Hannink, R. J. H. & Pascoe, R. T. *Nature* **258**, 703–705 (1975).
13. Claussen, N. & Rühle, M. in *Advances in Ceramics* Vol. 4 (eds Heuer, A. H. & Hobbs, L. W.) 137–163 (American Ceramic Society, Columbus, 1981).
14. Aksay, I. A., Lange, F. F. & Davis, B. I. *J. Am. ceram. Soc.* **66**, C190–C192 (1983).
15. Lange, F. F., Davis, B. I. & Aksay, I. A. *J. Am. ceram. Soc.* **66**, 407–408 (1983).
16. Lange, F. F. *J. Am. ceram. Soc.* **67**, 83–89 (1984).
17. Lange, F. F., Davis, B. I. & Wright, E. J. *Am. ceram. Soc.* **69**, 66–69 (1986).
18. Davidge, R. W. in *Fracture Mechanics of Ceramics* Vol. 2 (eds Bradt, R. C., Hasselman, D. P. H. & Lange, F. F.) 447–468 (Plenum, New York, 1973).
19. Alford, N. McN., Kendall, K., Clegg, W. J. & Birchall, J. D. *J. Am. ceram. Soc.* (submitted).
20. Ritter, J. E. & Davidge, R. W. *J. Am. ceram. Soc.* **67**, 432–437 (1984).
21. Weibull, W. *J. appl. Mech.* **18**, 293–297 (1951).
22. Kendall, K., Alford, N. McN. & Birchall, J. D. *Spec. Ceram.* **8**, 255–265 (1986).
23. Kendall, K. in *Tribology of Powders* (eds Briscoe, B. J. & Adams, M. J.) (Hilger, Bristol) in the press.
24. Eur. Pat. Publ. no. 0183453.
25. Br. Pat. Appl. no. 8709598.

Neodymium and strontium isotope content of microdiorite enclaves points to mantle input to granitoid production

P. Holden*†‡, A. N. Halliday†‡ & W. E. Stephens*

* Department of Geology, University of St Andrews, Purdie Building, St Andrews, Fife KY16 9ST, UK

† Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

Microdiorite enclaves are arguably the most poorly understood of the more prominent features of granitoids. Surprisingly, they have been invoked as support for completely different models for the origin of chemical zonation in granitoids, depending on whether they are interpreted as restites, syn-plutonic blobs of magma, remobilized cumulate or basified country-rock xenoliths. Here we present the first detailed neodymium and strontium isotope study of suites of such enclaves from two plutons. In both cases we find relatively high ϵ_{Nd} values, most readily explained if enclaves represent syn-plutonic injections of mafic magma into the host granitoid magma. As such they are the tangible remains of a very important process of mantle-derived heat and mass transfer operating in the generation of granitic magmas.

Microdioritic enclaves have been observed in intermediate

‡ Present address: Department of Geological Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA.

granitoids throughout the world¹⁻¹⁴. Typically, these enclaves are of spherical to discoid shape with a well-defined size range seldom exceeding 1 m in diameter, a feature probably related to the ascent velocity and/or yield strength of the magma¹⁴⁻¹⁶. Generally, the enclaves have igneous textures^{3,5-7}, are of fine to medium grain size¹⁻¹⁴ and are compositionally similar to quartz diorites¹⁻¹³. Enclaves have been variously interpreted as the residue after the partial digestion of country-rock fragments⁶, restite brought up from an igneous source region^{8,11}, remobilized basic precursors or cumulates (cognate enclaves)⁵⁻⁷, congealed syn-plutonic injections of basic liquid intruded into the granitic magma^{15,17-20}, and segregated immiscible liquids.

Recently, much emphasis has been placed on the role of enclaves as support for general models of granitoid generation. Chappell and co-workers⁸⁻¹¹ believe it is likely that more than 99% of mafic enclaves within granitoids of I-type in the Lachlan fold belt are restitic in origin⁸, such that their separation results in the differentiation of the magma producing linear 'unmixing' curves on Harker diagrams. Eichelberger¹⁷, on the other hand, believes that mafic enclaves represent the disrupted fragments of basic mantle-derived magmas which provide heat for crustal melting and granitoid genesis (see refs 18-20). Incorporation and mixing of such magmas with crustal melts would produce similar linear co-variation on Harker diagrams as would restite separation. A clear understanding of the origin of mafic enclaves is therefore essential to any hypothesis on the generation and compositional diversity of granitoids.

We chose to study the Criffell and Strontian plutons because their pluton geochemistries have already been detailed²¹⁻²⁷. Both plutons²⁸⁻³¹ (Fig. 1) are composite and normally zoned with abundant microdiorite enclaves within the outer, intermediate members but with relatively few in the evolved central portions. Country-rock fragments are rare in all members of both plutons.

Enclaves sampled from both plutons have igneous textures and thus at the outset appear to have been patches of more basic magma with oscillatory zoned plagioclase megacrysts and myrmekitic intergrowths providing primary magmatic features^{32,33}. Acicular apatites indicate rapid growth within a quench environment³⁴, but do not appear to be relict products of country-rock digestion³. Comprehensive textural descriptions for similar enclaves from around the world have been given elsewhere¹⁻⁷ and will be the subject of a more detailed publication.

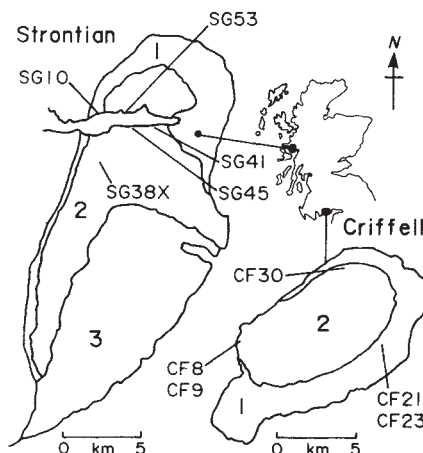


Fig. 1 Geological maps and sample locations at Strontian and Criffell. Samples were collected as host-enclave pairs immediately adjacent to each other. Suffix X represents enclaves; suffix G represents host granitoids. At Strontian: 1 = tonalite, 2 = granodiorite and 3 = biotite granite. At Criffell, 1 = clinopyroxene-hornblende-granodiorite, 2 = zoned hornblende-granodiorite to biotite granite.

Sm-Nd and Rb-Sr isotope results are given in Table 1 and Fig. 2. Initial isotopic ratios were calculated using age data published for Strontian³⁵ and Criffell²². Initial Nd isotope ratios indicate a lack of equilibration between host and enclave pairs, with the enclaves having consistently more radiogenic Nd (1 to 2ε units) for each complex. εNd values for the enclaves typically range from -0.7 to +1.6, whilst for host tonalites and granodiorites εNd range from -2.0 to +0.6. The εNd values of samples from the central parts of both plutons are less radiogenic than the outer tonalite-granodiorite members²¹, compatible with the larger budget of old crustal components in the more evolved portions indicated by Sr, O and U-Pb (zircon) studies^{22,23,36,37}. Initial Sr isotope results for host and enclave pairs (Figs 2 and 3) show little disequilibrium for both complexes, as has been reported elsewhere^{22,38,39}. Initial Sr isotopic ratios for whole-rock samples and host-enclave pairs from the tonalite-granodiorite members of both complexes^{22,38,39} are confined to a narrow range (0.7051-0.7061).

Table 1 Nd-Sr isotope data for Strontian and Criffell enclaves and immediately adjacent granitoid

Sample no.	Rb (p.p.m.)	Sr (p.p.m.)	⁸⁷ Rb/ ⁸⁶ Sr	⁸⁷ Sr/ ⁸⁶ Sr	(⁸⁷ Sr/ ⁸⁶ Sr) _i	Sm (p.p.m.)	Nd (p.p.m.)	¹⁴⁷ Sm/ ¹⁴⁴ Nd	¹⁴³ Nd/ ¹⁴⁴ Nd	(¹⁴³ Nd/ ¹⁴⁴ Nd) _i	εNd _i	T _{DM} (Gyr)
Strontian, t = 435 Myr												
SG10X	73	682	0.3082	0.70713 (3)	0.70523	8.00	63.4	0.1006	0.512446 (20)	0.51216	1.6	0.81
SG10G	51	1,260	0.1180	0.70606 (2)	0.70533	7.23	46.5	0.09314	0.512372 (20)	0.51211	0.6	0.90
SG38X	85	996	0.2480	0.70684 (4)	0.70530	5.96	34.7	0.1022	0.512426 (21)	0.51216	1.1	0.99
SG41X	100	1,397	0.2063	0.70656 (3)	0.70529	6.92	44.1	0.09503	0.512417 (20)	0.51215	1.3	0.81
SG41G	53	1,187	0.1284	0.70613 (4)	0.70534	7.08	44.3	0.09502	0.512309 (20)	0.51204	-0.8	0.95
SG45X	94	1,271	0.2134	0.70737 (3)	0.70605	1.71	37.8	0.09224	0.512364 (19)	0.51210	0.5	0.86
SG45G	50	1,235	0.1170	0.70535 (4)	0.70535	8.83	43.9	0.1213	0.512381 (19)	0.51204	-0.8	1.09
SG53X	111	971	0.3306	0.70706 (4)	0.70515	5.04	32.6	0.09359	0.512427 (21)	0.51216	1.6	0.79
SG53G	53	1,202	0.1271	0.70613 (4)	0.70534	5.42	32.8	0.09920	0.512382 (30)	0.51210	0.4	0.89
Criffell, t = 397 Myr												
CF8X	268	302	2.563	0.72087 (4)	0.70638	18.0	118.8	0.09160	0.512330 (20)	0.51209	-0.7	0.90
CF8G	160	572	0.8110	0.71059 (4)	0.70601	4.46	26.3	0.1015	0.512288 (20)	0.51202	-2.0	1.03
CF9X	213	454	1.361	0.71346 (3)	0.70576	6.74	44.5	0.09154	0.512375 (21)	0.51214	0.1	0.84
CF9G	150	475	0.9116	0.71104 (4)	0.70588	3.45	25.2	0.08051	0.512311 (21)	0.51210	-0.5	0.84
CF21X	167	882	0.5494	0.70829 (4)	0.70518	7.52	50.3	0.09030	0.512340 (20)	0.51211	-0.4	0.87
CF21G	128	884	0.4189	0.70775 (4)	0.70538	3.75	25.8	0.08787	0.512312 (20)	0.51208	-0.8	0.89
CF23X	108	759	0.4108	0.70778 (4)	0.70546	18.6	145.9	0.09399	0.512386 (21)	0.51214	0.3	0.84
CF23G	105	849	0.3601	0.70741 (3)	0.70537	5.36	36.0	0.09096	0.512314 (19)	0.51208	-1.0	0.91
CF30X	368	354	3.014n	0.72295 (4)	0.70591	15.41	107.8	0.09642	0.512346 (21)	0.51212	-0.1	0.84
CF30G	150	539	0.8062	0.71065 (3)	0.70601	4.60	28.7	0.09619	0.512328 (21)	0.51208	-0.9	0.93

Nd isotope analyses were carried out on a VG Isomass 54E mass spectrometer, and Sr isotopic compositions were determined on a VG Micromass MM30B machine. Ion exchange techniques were developed from those of Richard *et al.*⁴³. ¹⁴³Nd/¹⁴⁴Nd and ⁸⁷Sr/⁸⁶Sr are normalized to ¹⁴⁶Nd/¹⁴⁴Nd = 0.7219 and ⁸⁶Sr/⁸⁶Sr = 8.37521, respectively. Bulk-Earth parameters used to calculate εNd values and T_{DM} ages⁵¹ are ¹⁴³Nd/¹⁴⁴Nd = 0.512638, ¹⁴⁷Sm/¹⁴⁴Nd = 0.1966. Numbers in parentheses represent two σ errors on the mean. All Rb, Sr, Sm and Nd concentrations determined by isotope dilutions.

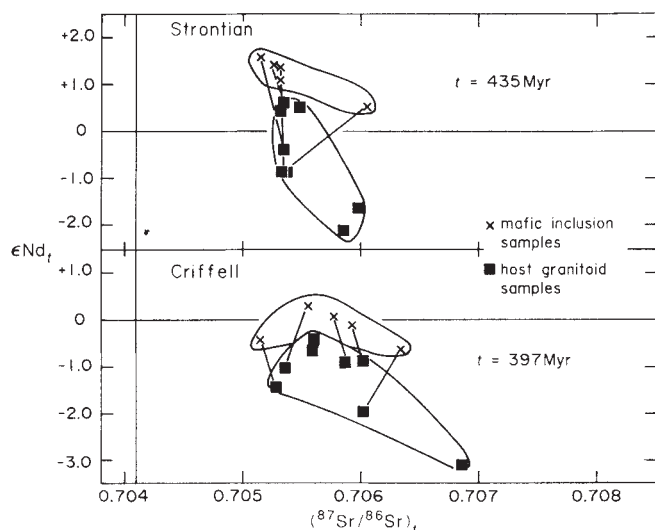


Fig. 2 ϵNd_t against $(^{87}\text{Sr}/^{86}\text{Sr})_t$ (where $t = 435$ Myr for Strontian³⁸ and 397 Myr for Criffell²²). Additional data from refs 22, 37, 38. Tie-lines join immediately adjacent host-enclave pairs.

It is apparent from Fig. 2 that tie-lines joining host and enclave pairs have no consistent slope, indicating that it is unlikely that the host equilibrated with enclaves derived from a single homogeneous source, as such homogenization, would be expected to produce a suite of similar tie-lines. Furthermore, isotopic data for nodules of pyroxenite, spinel lherzolite and basic granulite⁴⁰⁻⁴³ erupted during the Carboniferous suggest that the mantle and mafic lower crust as (apparently) sampled nearest to Strontian and Criffell has ϵNd values of 0 to +4 but variable $^{87}\text{Sr}/^{86}\text{Sr}$ (refs 40-43). The nodule data encompass the field of granulite enclaves reported herein and therefore little equilibration with the host granitoid is necessary to explain the isotopic data. However, note that grain boundary re-adjustment of the major repository of Sr (plagioclase) has afforded ample opportunity for at least partial late equilibration, whilst the repositories of Nd, the accessory phases (particularly zircon and apatite), show no evidence of recrystallization. Thus, whilst Sr may well have equilibrated, Nd equilibration was most probably minimal, a finding consistent with experimental data on rates of diffusion^{43,44}.

In the two plutons investigated the fact that initial Nd isotopic ratios in enclaves are higher than their hosts by 1 to 2 ϵ units rules out a cogenetic origin for these enclaves unless the host has suffered considerable contamination from material with low ϵNd_t . This appears to be unlikely as zircon separates from the outer portions of the Strontian pluton do not contain a Pb isotopic memory³⁵, precluding major assimilation of local Precambrian country rock or inferred older basement. Furthermore, disrupted chilled margins or spontaneously segregated immiscible melts are not tenable enclave sources as these would also be expected to have the same Nd isotopic composition as the host magma. It is therefore incorrect to describe such enclaves as 'cognate' or as 'autoliths'. Since Sr approaches isotopic equilibrium between host and enclave, any hypothesis on the origin of enclaves using Sr isotopic studies alone must be viewed with scepticism.

Figure 3 shows clearly that digested local country-rock fragments can also be discounted as viable sources for microdioritic enclaves at Strontian. Equilibration with country rocks⁴⁵ (such as Moine metasediments) cannot produce residual enclaves with ϵNd values greater than Bulk Earth at the time of emplacement. The same is true for the deep crust if it had $\epsilon\text{Nd} < 0$. Similarly, the Southern Uplands sediments are not a viable source for Criffell enclaves since all those analysed (with the exception of one basic-clast greywacke²² from the opposite end of the accretionary prism) have $\epsilon\text{Nd} < 0$ (refs 21, 46). Oscillatory zoned

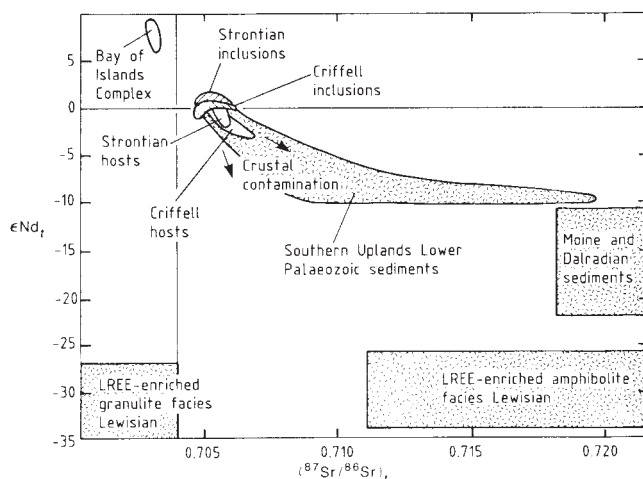


Fig. 3 ϵNd_t against $(^{87}\text{Sr}/^{86}\text{Sr})_t$ for Strontian, Criffell and local crustal rocks (refs 22, 47, 48, 49 and P.H., unpublished data).

plagioclase megacrysts within the enclaves provide clear evidence for an original magmatic origin^{32,33}, whilst acicular apatites suggest that the enclaves were the product of quenching⁴⁰ of an evolving magma more basic than the host, possibly at high level (but obviously not at the final 'resting' position exposed within the pluton). Nd isotopic data support this hypothesis and are consistent with Nd isotopic data for those local nodule suites which are thought to be mantle-derived⁴⁰⁻⁴⁴.

The implication of this study is that magmas clearly derived from the mantle were intruded into the crust in close association with granitoid production. This is not thought to be coincidental and it is difficult to escape the conclusion that a major input of melt and heat from the mantle has substantially aided crustal melting and granitoid genesis during the late Caledonian in Scotland. Furthermore, the identification of a substantial bulk mantle contribution to the isotopic composition of granitoids indicates the continuing input of mass from the mantle to the crust. This is in marked contrast to the explanations given for the granites in the Lachlan fold belt (south-east Australia), in which granitoid production is seen purely in terms of intracrustal melting during period of high heat flow⁸⁻¹¹.

Enclaves are a ubiquitous feature of I-type granitoids worldwide and the identification of enclaves as mantle-derived with pristine or near-pristine Nd isotopic compositions opens up new possibilities for the study of mantle and mantle-crust relationships, in terms of melt dynamics and crustal growth. As such, enclaves should form an integral part of any study into the origins and compositional diversity of intermediate to acid magma, but the usefulness of Sr and Pb isotopic systematics to such studies is severely limited by their ready equilibration. Nd isotopic data is essential if the correct interpretation is to be made.

That the microdiorite enclaves plot at the end of the Nd-Sr isotope array for the host granitoids in Fig. 2 is not regarded as coincidental and these enclaves give the clearest indication to date of the relative importance of mantle-derived Nd to the overall Nd budget of the granitoids. The importance of mantle-derived components in the other granitoids cannot be adequately assessed until similar Nd isotopic data is made available.

This project was supported by the NERC. The Isotope Geology Unit at SURRC is also supported by NERC and the Scottish Universities. We thank Jo Hamilton for some repeat analyses and J. Hutchinson and F. MacLaren for laboratory assistance.

Received 15 June; accepted 4 September 1987.

1. Didier, J. *Granites and Their Enclaves* (Elsevier, Amsterdam, 1983).
2. Cantagrel, J. M., Didier, J. & Gourgou, A. *Phys. Earth planet. Interiors* **35**, 63-76 (1984).
3. Vernon, R. H. *Nature* **309**, 438-439 (1984); *J. Proc. R. Soc. NSW* **116**, 77-103 (1983).
4. Reid, J. B., Evans, O. C. & Fates, D. C. *Earth planet. Sci. Lett.* **66**, 243-261 (1983).
5. Pabst, A. *Univ. Calif. Publ. Geol. Sci.* **17**, 325-386 (1928).

6. Grout, F. F. *Bull. Geol. Soc. Am.* **48**, 1521-1572 (1937).
7. Ferstata, G. B. & Borodina, N. S. *Inst. Geol. Rev.* **19**, 458-468 (1977).
8. White, A. J. R. & Chappell, B. W. *Tectonophysics* **43**, 7-22 (1977).
9. Chappell, B. W. *Phil. Trans. R. Soc. A310*, 693-707 (1984).
10. Compston, W. & Chappell, B. W. in *The Earth: Its Origin, Structure and Evolution* (ed. McElhinny, M. W.) 377-426 (Academic, London, 1979).
11. McCulloch, M. T. & Chappell, B. W. *Earth planet. Sci. Lett.* **58**, 51-58 (1982).
12. Whalen, J. B. *J. Petrol.* **26**, 603-632 (1985).
13. Bender, J. F., Hanson, G. N. & Bence, A. E. *Earth planet. Sci. Lett.* **58**, 330-344 (1982).
14. Oxburgh, E. R. & McRae, T. *Phil. Trans. R. Soc. A310*, 457-472 (1984).
15. Sparks, R. S. J., Sigurdson, H. & Wilson, L. *Earth planet. Sci. Lett.* **35**, 234-238 (1977).
16. Spera, F. J. in *Physics of Magmatic Processes* (ed. Hargreaves, R. B.) (Princeton University Press, 1980).
17. Eichelberger, J. C. *Nature* **275**, 21-27 (1978); **288**, 446-450 (1980).
18. Grey, C. M. *Earth planet. Sci. Lett.* **70**, 47-60 (1984).
19. Domenick, M. A., Kistler, R. W., Dodge, F. C. W. & Tatsumoto, M. *Bull. Geol. Soc. Am.* **94**, 713-719 (1983).
20. Hildreth, W. *J. geophys. Res.* **86**, 10153-10192 (1981).
21. Halliday, A. N. *Nature* **307**, 229-233 (1984).
22. Halliday, A. N., Stephens, W. E. & Harmon, R. S. *J. Geol. Soc.* **137**, 329-348 (1980).
23. Halliday, A. N. & Stephens, W. E. *Phys. Earth planet. Interiors* **35**, 89-109 (1984).
24. Stephens, W. E. & Halliday, A. N. *Phil. Trans. R. Soc., Edin.* **75**, 259-273 (1984).
25. Harmon, R. S., Halliday, A. N., Clayburn, J. A. P. & Stephens, W. E. *Phil. Trans. R. Soc. A310*, 709-742 (1984).
26. Thirwall, M. T. *Earth planet. Sci. Lett.* **58**, 27-50 (1982).
27. Pidgeon, R. T. & Aftalion, M. *Geol. J.* **10**, 183-220 (1979).
28. Sabine, P. A. *Bull. geol. Surv.* **20**, 6-42 (1962).
29. MacGregor, A. G. & Kennedy, W. Q. *Sum. Proc. Geol. Surv.* (1931) 105 (1932).
30. Phillips, W. J. *J. geol. Soc.* **112**, 221-240 (1956).
31. Phillips, W. J., Fuge, R. & Phillips, N. J. *J. geol. Soc.* **138**, 351-366 (1981).
32. Vance, J. A. *Am. J. Sci.* **260**, 746-760 (1962).
33. Hibbard, M. J. *Contrib. Miner. Petrol.* **90**, 1047-1062 (1979); *Geol. Soc. Am. Bull.* **76**, 158-170 (1981).
34. Wyllie, P. J., Cox, K. G. & Biggar, G. M. *J. Petrol.* **3**, 238-242 (1962).
35. Halliday, A. N., Aftalion, M., van Breemen, O. & Jocelyn, J. in *Caledonides of the British Isles Reviewed* (eds Harris, A. L., Holland, C. H. & Leake, B. E.) 653-661 (Scottish Academic Press, 1979).
36. Halliday, A. N. in *Melting Migmatites and Metamorphism*, 54-61 (Shiva, Nantwich, 1983).
37. Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287**, 279-284 (1980).
38. Pankhurst, R. J. in *Origin of Granite Batholiths*, 18-33 (Shiva, Nantwich, 1979).
39. Stephens, W. E. & Halliday, A. N. *Bull. geol. Soc. Am.* **91**, 165-170 (1980).
40. Halliday, A. N. *et al. Scott. J. Geol.* **21**, 465-491 (1985).
41. Halliday, A. N., Menzies, M. A., Palacz, Z., Hunter, R. H. & Hawkesworth, C. J. *Trans. Am. geophys. Un.* **66**, 414 (1985).
42. Menzies, M. A. & Halliday, A. N. *Trans. R. Soc. Edin.: Earth Sci.* **75**, 298-299 (1984); *27th Int. Geol. Congr. Abstr.* **5**, 348 (1984).
43. Margaritz, M. & Hofmann, A. W. *Geochim. cosmochim. Acta* **42**, 595-605 (1978).
44. Margaritz, M. & Hofmann, A. W. *Geochim. cosmochim. Acta* **42**, 847-858 (1978).
45. Frost, C. D. & O'Nions, R. K. *J. Petrol.* **26**, 515-544 (1985).
46. O'Nions, R. K., Hamilton, P. J. & Hooker, P. J. *Earth planet. Sci. Lett.* **63**, 229-240 (1983).
47. Farmer, G. L. & DePaolo, D. J. *J. geophys. Res.* **88**, 3379-3401 (1982).
48. Hamilton, P. J., Evenson, N. M., O'Nions, R. K. & Tarney, J. *Nature* **277**, 25-28 (1979).
49. Jacobson, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7249-7445 (1979).

Magic-angle-spinning NMR shows the aluminosilicate framework of ultramarine to be disordered

Jacek Klinowski*, Stuart W. Carr*, Stephen E. Tarling† & Paul Barnes†

* Department of Physical Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EP, UK

† Industrial Materials Group, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK

²⁹Si magic-angle-spinning NMR (MAS NMR) spectra of synthetic ultramarine show that the structure of its aluminosilicate framework is different from that of lazurite, its natural counterpart. Si-O-Si, Si-O-Al and Al-O-Al linkages are simultaneously present and Si and Al sites are disordered. This agrees with the results from the Rietveld refinement of neutron diffraction data and suggests that either the mechanisms of formation of lazurite and ultramarine are different or that the ordering process occurs on a long time scale.

Ultramarine is the synthetic equivalent of the natural aluminosilicate lazurite. Brought from Afghanistan to Europe by Marco Polo in the 1270s, lazurite became, on account of its wonderfully intense colour which is unaffected by sunlight, the source of the royal blue pigment much favoured by Renaissance painters. For centuries it was more valuable than gold; in 1828 it was first prepared synthetically (by furnacing a mixture of kaolin, sodium carbonate and sulphur) in response to a competition for a commercial process^{1,2}.

The chemical formula of ultramarine is ideally



and the framework structure is that of zeolite sodalite³⁻⁷ (see Fig. 1). In blue ultramarine the chromophore groups are the strongly paramagnetic S_3^- and S_2^- anions enclathrated inside the sodalite cages⁸⁻¹⁰. Other colours can be prepared by ion exchange and by chemical modification of the chromophore group.

Our task is to establish the precise nature of the distribution of Si and Al among the tetrahedral framework sites in ultramarine. In the case of its natural equivalent, the answer is known from single-crystal diffraction¹¹: lazurite is cubic, space group $P4_3n$ with $a_0 = 9.1 \text{ \AA}$. The framework is fully ordered and consists of alternating SiO_4^{4-} and AlO_4^{5-} tetrahedra. Ultramarine,

however, is microcrystalline and furthermore contains four ten-electron species (O^{2-} , Na^+ , Al^{3+} and Si^{4+}) with very similar scattering for X-rays. We have therefore resorted to high-resolution solid-state ²⁹Si and ²⁷Al NMR which readily monitors the type, quantity and distribution of Si and Al sites, irrespective of crystal size¹².

We have studied two ultramarines: standard blue (sample 1) and pink (sample 2) prepared from sample 1 by thermal treatment in the presence of chlorine. Both samples were analysed by wet chemical methods and by X-ray fluorescence and were found to have the same Si and Al contents: Si/Al = 1.11 ± 0.03 . Bromine-soluble amorphous impurities amounted to 5 weight%; as these are derived from kaolinite (with Si/Al = 1.00) no correction for the composition of the crystalline part of the sample was necessary. Both ultramarines were checked by X-ray and neutron diffraction and found to have the same structure. ESR showed that sample 1 was strongly and sample 2 slightly paramagnetic.

²⁹Si MAS NMR spectra of aluminosilicates are composed of up to five Si(*n*Al) signals where *n*, which can be 0, 1, 2, 3 and 4, is the number of Al atoms linked through bridging oxygens

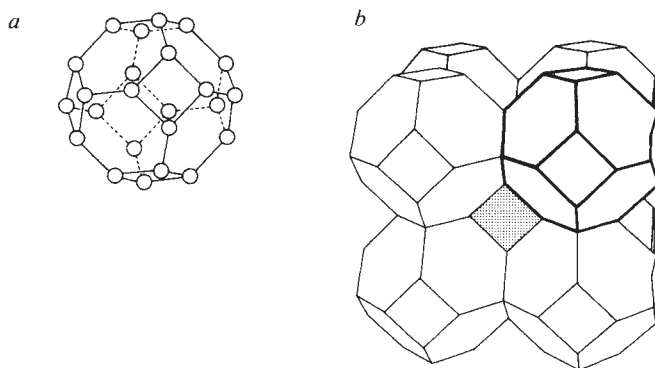


Fig. 1 The aluminosilicate framework of lazurite/ultramarine. *a*, The 14-hedral sodalite cage (also termed cubooctahedron) is built of eight 6-membered and six 4-membered rings. Tetrahedral silicon and aluminium atoms, represented by circles, are linked through oxygen atoms (not shown, for clarity). *b*, The sodalite cage (marked in bold lines) is a space-filling solid. The framework is formed by direct face-sharing of 4-membered rings in the neighbouring sodalite cages and is found in the minerals of the sodalite family (sodalite, nosean, hauyne and lazurite). Sodalite cages normally contain guest molecules or ions (not shown).

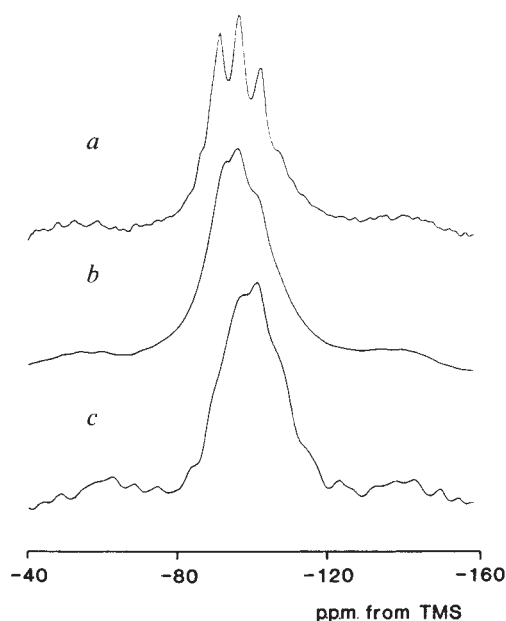


Fig. 2 ^{29}Si MAS NMR spectra of ultramarines. *a*, Pink ultramarine (sample 2); *b*, the spectrum in (*a*) with 260 Hz line-broadening; *c*, standard blue ultramarine (sample 1). The spectra were obtained at 79.5 MHz using a Bruker MSL-400 multinuclear spectrometer with 45° resonant pulses and 15 s recycle delay. Samples were spun in air at 4 kHz.

to the central silicon^{12,13}. The spectrum of sample 2 given in Fig. 2*a* and Fig. 3 indeed contains five signals: a shoulder at -87.8 p.p.m. from tetramethylsilane (TMS) which we assign to Si(4Al) units, and lines at -92.5, -97.6, -103.4 and -108.2 p.p.m. which we assign to Si(3Al), Si(2Al), Si(1Al) and Si(0Al), respectively. This assignment is supported by earlier work: the neighbouring Si(*n*Al) signals are about 5 p.p.m. apart¹³, and the single Si(4Al) line in synthetic sodalite comes at -86.5 p.p.m. (J.K., unpublished observation).

The spectrum of sample 1 (Fig. 2*c*) is much less well-resolved than that of sample 2: there is a broad line centred at about -100 p.p.m. with some indication of fine structure. The reason is paramagnetic broadening: by applying 260 Hz exponential line-broadening to the spectrum of sample 2 we obtain the spectrum in Fig. 2*b*, which is very similar to that of sample 1 (Fig. 2*c*), but the latter is paramagnetically shifted by about 5 p.p.m. to high field. It is thus clear that, despite different resolution, the intrinsic line intensities of samples 1 and 2 are the same.

Deconvolution of the spectrum of sample 2 using gaussian line shapes (Fig. 3) gives the ratio of intensities of the lines

Table 1 Spectral intensities

Signal	Measured	Perfectly random	Restricted random
Si(4Al)	7.3	5.0	65.9
Si(3Al)	28.6	22.4	29.0
Si(2Al)	30.2	37.3	4.8
Si(1Al)	20.1	27.6	0.3
Si(0Al)	13.8	7.7	0

These values are measured and calculated (from the binomial formula) spectral intensities, normalized to a total of 100, of the Si(*n*Al) signals in ^{29}Si MAS spectra of ultramarine (Si/Al=1.11) either for a perfectly random Si,Al distribution or for a random distribution of Si and Al, but one in which Al-O-Al linkages are forbidden.

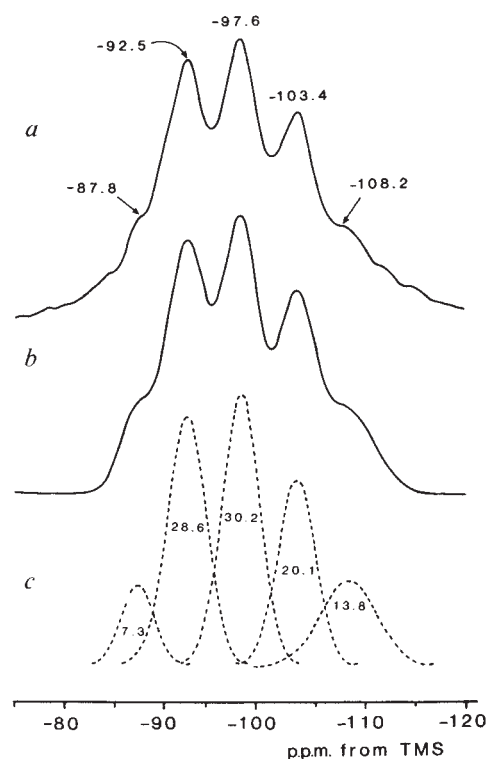


Fig. 3 *a*, The ^{29}Si MAS NMR spectrum of pink ultramarine; *b*, can be simulated; *c*, using five gaussian peaks with the ratio of intensity as indicated.

as Si(4Al):Si(3Al):Si(2Al):Si(1Al):Si(0Al) = 7.3:28.6:30.2:20.1:13.8. This result has immediate implications for the Si,Al distribution in the framework. We shall consider two possibilities:

(1) *The distribution of Si and Al is random*, but Al-O-Al linkages are not allowed. This restriction¹⁴ follows from Pauling's electrostatic valence principle and applies to sodalite, which is isostructural with ultramarine. The probability for an Si-O-Al linkage to occur is then $p = 1/r$, where r is the Si/Al ratio. The probability of an Si-O-Si linkage is of course $1 - p$. For $r = 1$ we have $p = 1$ and all linkages must be of the Si-O-Al type, that is, only Si(4Al) units can be present. Thus for $r = 1$, the "restricted" model requires perfect ordering of Si and Al. The ^{29}Si NMR spectrum consists of a single line as observed for sodalite. As r increases, other units appear with relative populations given by the binomial formula as p^4 for Si(4Al), $4(1-p)p^3$ for Si(3Al), $6(1-p)^2p^2$ for Si(2Al), $4(1-p)p$ for Si(1Al) and $(1-p)^4$ for Si(0Al). Table 1 gives these populations for our samples. It is clear that the number and relative intensities of the measured NMR signals are quite incompatible with the restricted model.

(2) *The distribution of Si and Al is truly random* in that Al-O-Al linkages are allowed. In this case $p = 1/(r+1)$ and the ideal sample with $r = 1$ contains all five types of Si(*n*Al) units in the population ratio of 6.25:25:37.5:25:6.25. Table 1 shows that the values calculated for our samples ($r = 1.11$) are in good agreement with experiment. Simulation based on calculated line intensities produces a spectrum which is very similar to that in Figs 2*a* and 3.

Thus ^{29}Si MAS NMR unambiguously demonstrates that the distribution of Si and Al in ultramarine is fully random. This is in striking agreement with the Rietveld refinement of neutron powder diffraction data¹⁵. Refinement in space group P4₃n, which provides for alternating Si and Al sites, leads to R -factors $> 20\%$ whereas refinement in the fully random space group I4₃m gave R -factors below 15%.

Al-O-Al linkages are not excluded on theoretical grounds¹⁶

and are indeed found in, for example, the entire family of sodalites with frameworks composed exclusively of AlO_4^{3-} tetrahedra¹⁷⁻¹⁹. These, like the ultramarines, are prepared pyrolytically. Hydrothermally synthesised zeolites, including sodalite, never contain Al-O-Al linkages¹² and there is evidence²⁰ that they are forbidden even in aluminosilicate anions in solution.

²⁷Al MAS NMR spectra of our samples (not shown) each consist of a single signal typical of tetrahedrally coordinated Al. The linewidth of the signal from blue ultramarine (1,800 Hz) is much greater than that from the pink sample (840 Hz). But the latter linewidth is still double that of synthetic sodalite,

which is partly due to the residual paramagnetism, and partly to the disordered environment of Al in ultramarine. In hydrothermally prepared zeolites all framework Al atoms are in an identical Al(4Si) environment, whereas in ultramarine all five Al(*n*Si) environments are present (*n* = 0, 1, 2, 3 and 4).

We conclude that ultramarine is the disordered form of lazurite. It appears that either the mechanisms of formation of the two species are different or that the ordering process in the disordered precursor proceeds on a very long time scale. This is known to occur among feldspars.

We are grateful to Reckitt's Colours Ltd and the SERC for a CASE studentship to S.E.T.

Received 18 June; accepted 16 September 1987.

- Guimet, J. B. *Ann. Chim.* **46**, 431-2 (1831).
- Gmelin, C. G. *Jahr. Chem. Phys. (Halle)* **24**, 360-380 (1831).
- Jaeger, F. M. *Trans. Faraday Soc.* **25**, 320-345 (1929).
- Gruner, E. Z. *Anorg. Allg. Chem.* **224**, 351-368 (1935).
- Leschewski, K. *Angew. Chem.* **48**, 533-6 (1935).
- Podschus, E., Hofmann, U. & Leschewski, K. Z. *Anorg. Allg. Chem.* **228**, 305-333 (1936).
- Pauling, L. Z. *Krist.* **74**, 213-225 (1930).
- Clark, R. J. H. & Franks, M. L. *Chem. phys. Lett.* **34**, 69-72 (1975).
- Cotton, F. A., Harmon, J. B. & Hedges, R. M. J. *Am. chem. Soc.* **98**, 1417-1424 (1976).
- Clark, R. J. H. & Cobbold, D. G. *Inorg. Chem.* **17**, 3169 (1978).

- Hassan, I., Peterson, R. C. & Grundy, H. D. *Acta crystallogr.* **C41**, 827-832 (1985).
- Klinowski, J. *Progress in NMR Spectroscopy* **16**, 237-309 (1984).
- Lippmaa, E., Mägi, M., Samoson, A., Engelhardt, G. & Grimmer, A.-R. J. *Am. chem. Soc.* **102**, 4889-4893 (1980).
- Loewenstein, W. *Am. Miner.* **39**, 92-96 (1953).
- Tarling, S. E., Barnes, P. & Klinowski, J. *Acta crystallogr.* (in the press).
- Barrer, R. M. & Klinowski, J. *Phil. Trans. R. Soc. A285* (No. 1329), 637-676 (1977).
- Saalfeld, H. & Depmeier, W. *Krist. und Technik* **7**, 229-233 (1972).
- Depmeier, W. J. *appl. Crystallogr.* **12**, 623-626 (1979).
- Kondo, R. *Yogyo Kyokaishi Shi* **71**, 1-8 (1965).
- Müller, D., Hoebbel, D. & Gessner, W. *Chem. phys. Lett.* **84**, 25-29 (1981).

Isolating individual chains of selenium by incorporation into the channels of a zeolite

O. Terasaki*, K. Yamazaki*, J. M. Thomas†||, T. Ohsuna*, D. Watanabe*, J. V. Sanders‡ & J. C. Barry§

* Department of Physics, Tohoku University, Sendai 980, Japan

† Davy Faraday Research Laboratory, The Royal Institution, 21 Albemarle Street, London W1X 4BS, UK

‡ Division of Materials Science, CSIRO, Normanby Road, Clayton, Victoria 3168, Australia

§ Center for Solid State Science, Arizona State University, Tempe, Arizona 85287, USA

One of the disappointments to emerge from the recent intensive study of low-dimensional solids has been the increasing realization that relatively few of the investigated systems exhibit genuine one- or two-dimensional behaviour. Measured electronic and optical properties of crystalline and amorphous low-dimensional solids signify that there is significant interchain or interlayer perturbation. One possible way of eliminating such perturbation in so-called one-dimensional systems is to accommodate individual chains within the pores of a host dielectric material, which, preferentially, should itself be crystalline and well ordered. Here we report the use of combined high-resolution electron microscopy and computer simulation, to achieve the successful incorporation of chains of selenium into a synthetic mordenite, a zeolite which has one-dimensional channels (diameter ~ 7 Å) running parallel to its *c*-axis. Remarkably, the uptake of selenium, which converts the white mordenite into an orange colour, occurs in a patchwise fashion leading to domains of occupied channels. The incorporation is readily accomplished thermally, and can be effected with a range of other zeolitic hosts, including ZSM-5, ZSM-23 and zeolites L and Y.

Zeolites, which are open structures made up of corner-sharing SiO_4^{4-} and AlO_4^{5-} tetrahedra, are attractive as potential hosts for a variety of linear and cluster-like guests possessing novel electronic properties. And it has already been demonstrated by Bogomolov *et al.*¹ that the interconnected cavities (diameter ~ 13 Å) of faujasitic zeolites (for example, Mg^{2+} and Na^{+} -

exchanged zeolite X) can accommodate clusters of 16 or 23 atoms of Te. Moreover, zeolite A, with rather smaller cages and substantially smaller aperture diameters than zeolites X and Y, 4 Å as against 7.4 Å, can also accommodate tetranuclear clusters of alkali metals²⁻⁴ (for example, Na_4^{3+} and K_4^{3+}), as well as larger metallic clusters⁵ without severe disruption of the parent aluminosilicate framework. It has been reported that the one-dimensional tunnel structure of dehydrated mordenite may serve as a suitable receptacle for Te⁶ and Se^{6,7}, and that their optical properties are different from those of bulk crystals. But there has been no direct evidence that Se chains sit in the main channels of mordenite. We have prepared crystalline samples of mordenite in which approximately three quarters of the main channels and none of the small ones (diameters ~ 7.0 and 3.5 Å, respectively) are filled with extended chains of Se. Direct proof of the incorporation comes from high-resolution electron microscopy. The centres of the main channels in mordenite are separated from one another by ~ 13.5 Å.

Synthetic mordenite (idealized formula in the Na^{+} ion-exchanged form: $\text{Na}_8\text{Al}_8\text{Si}_{40}\text{O}_{96}\cdot 24\text{H}_2\text{O}$, $a = 1.81$, $b = 20.5$ and $c = 7.5$ Å) supplied by the Catalysis Society of Japan (JRC-Z-HM20) was used in its H^{+} form. It was first dehydrated by heating to 350 °C and then kept in a vacuum of better than 10^{-5} torr. This material could be loaded with Se when an appropriate quantity of the latter (99.999% purity) was heated with the mordenite in a sealed Pyrex tube for 3 h at 400 °C and then cooled to room temperature. After this heat treatment mordenite changed colour from white to orange. As can be seen from the colour, the diffuse-reflection spectrum is different from that of trigonal-Se and of amorphous-Se⁸. Their colours are closer to red than to orange. Fuller results will be published elsewhere. The specimen was crushed in an agate mortar under acetone and collected on a microgrid for electron microscopic study. Samples were examined in an electron microscope (JEM-1000) fitted with a top entry goniometer. The Se, Si and Al contents were determined for many crystals from the intensity of the X-ray L and K lines stimulated by 20-kV electrons when samples were examined in a scanning electron microscope (JSM-T330). Because Al K-lines overlap those of the Se L-lines, the contribution of Se-L was derived from its known relationship to Se-K emission.

Simulated electron microscopic images for the parent H^{+} form of mordenite and mordenite with Se in the main channels were computed using the procedures described elsewhere⁹⁻¹¹. For the 1-MeV microscope the instrumental characteristics were: Cs, 11 mm; beam divergence, 0.5 mrad; spread of focus 260 Å and

|| To whom correspondence should be addressed.

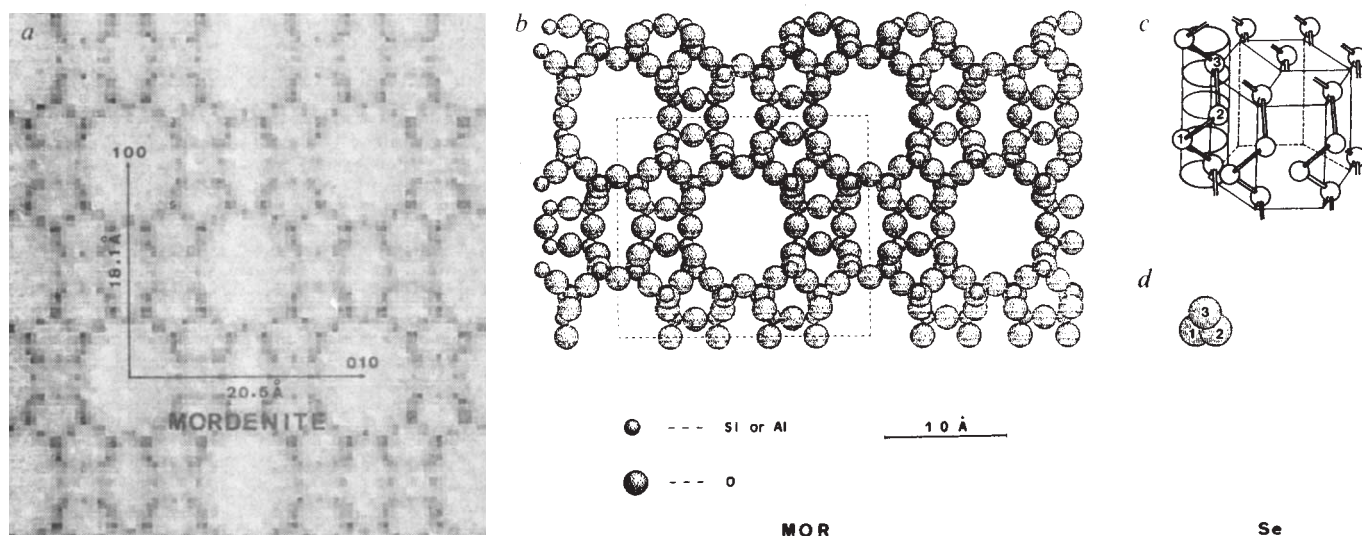


Fig. 1 *a*, Projected potential density (along [001]) of mordenite and *b*, schematic drawing of its framework structure. The structure of crystalline Se (so called A8 type with $a = 4.37$, $c = 4.95$ Å) is shown, in elevation, in *c*; and a projection of the chain of Se along the c -axis is shown in *d*. Note that, from the scale of these drawings, only one chain of Se can be accommodated in the main channels of mordenite.

resolution 1.9 Å. The atomic coordinates for the framework elements of the mordenite were taken from X-ray data¹²; the positions of the Se atoms in the incorporated chains within mordenite were taken to be the same as those for crystalline Se. Only one chain, (see Fig. 1) is capable of being accommodated in the large channels of mordenite, the projected potential density of which, in the absence of Se, is shown in Fig. 1*a*. Full details of the computed images for both the empty and occupied channels, as well as the corresponding diffraction patterns will be published elsewhere. Figures 2 and 3 are high-resolution micrographs of the Se-incorporated mordenite. Both transmission micrographs differ sharply from those corresponding to the parent mordenite, the images of which, show unoccupied channels that are visible from their clear, white contrast.

In Fig. 2 black and white contrast appears as wavy bands, which are roughly perpendicular to the b -axis, there being a periodicity of ~ 100 Å. There is no doubt that Se causes the

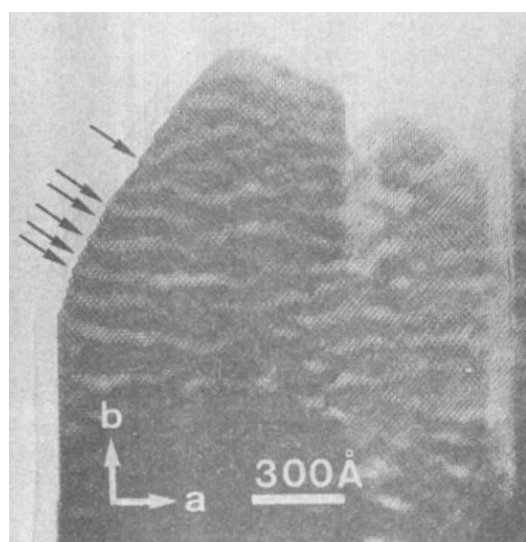


Fig. 2 High-resolution electron micrograph of mordenite into which Se has been inserted. Note the subunit cell steps at the periphery (see text).

contrast. The band contrast does not reverse when the value of defocus changes from over to under focus. The area corresponding to the black contrast is approximately three quarters of the total. This fraction is not dependent upon crystal thickness, even though the relative contrast is. This black contrast, which shows up in samples containing Se, but not in those devoid of it, cannot arise from surface segregated, crystalline Se, as the diffraction pattern of the latter is absent. At even higher magnifications (Fig. 3), the images in the white bands are indistinguishable from those of mordenite containing no Se, that is, the main and subsidiary channels are imaged as white dots. In the dark bands, however, the contrast of the main channels is black, whereas that of the small channels remains white. This information, along with computed images confirms that the Se is, indeed, incorporated into the large channels of mordenite. There is one further confirmatory fact. We may estimate the ratio of Se to tetrahedral (Si + Al) atoms from the structures assumed before. (Note that the unit cell of mordenite contains 48 tetrahedrally bonded atoms (Si + Al) and that it has two channels (total length 2×7.5 Å). Nine Se atoms can, therefore, be accommodated in the unit cell.) We may also derive this ratio from the intensity of the X-ray emission. These estimates come out to be 0.19 (that is, $9/48$) and 0.15 respectively; and the inverse ratio of these (0.8) is in quite close agreement with the estimated ratio of dark-to-light contrast in Figs 2 and 3.

It is not yet certain what gives rise to this patchwise, modulated (dark-light) nature of the Se-containing mordenite. It cannot be attributed to undersaturation, for such modulation between dark and light regions in the image is found even for specimens prepared under conditions when Se was present in excess. Several possibilities are likely. First, we could be seeing here a direct consequence of the well-known tendency for mordenite to exhibit either narrow-pore or wide-pore behaviour². If the mordenite contains co-existent narrow and wide pores, might it not be that the white regions correspond to the narrow pores that show little aptitude to take up Se or any other guest? Second, could not the differences observed arise from local (patchlike) variation in Si/Al ratio? Perhaps an Al-rich region has constrained pores that are less likely to take up Se. Third, we may be seeing in the dark-light contrasts, the manifestations of spinodal decomposition. Lastly, note that the dark regions all terminate with stepped edges (arrowed in Fig. 2), so that it is conceivable for the sticking coefficients of the Se as it impinges upon the mordenite to be very different depending upon the

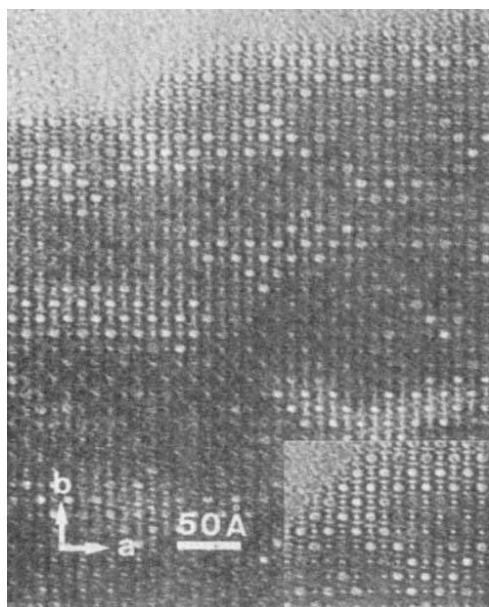


Fig. 3 At yet higher magnification (compare Fig. 2), the regions of light contrast are seen to be indistinguishable from mordenite devoid of any guest species. At the regions of dark contrast, caused by uptake of Se (see text) the small channels are still empty as indicated by the small white dots.

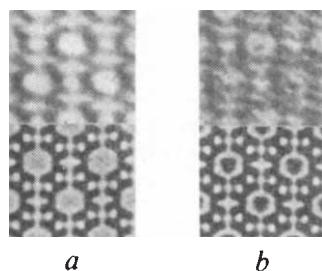


Fig. 4 The upper images are taken from white (a) and black (b) regions. The lower ones are simulated images for virgin mordenite (a) and mordenite with Se (b) at the conditions of specimen thickness of 75 Å and 950-Å defocus.

precise topography of the zeolite that it encounters. No patch-wise contrast was observed in zeolite L containing Se prepared using the method described here.

Until more work (now under way) is done, it is premature to decide which of the explanations is more likely to be correct. Of immediate practical importance, however, is the fact that Se chains can be readily incorporated into a number of distinct types of zeolites. There are very many zeolites which have channels of different aperture and cavity size, and different manner of channel connections, so it should, in future, be possible to arrange for extended chains of Se or other guests to be incorporated into dielectric hosts where the degree of perturbation between 'isolated' chains can be delicately controlled. Our preliminary studies reveal that the diffuse reflectance spectra of Se are different as between the zeolitic hosts mordenite, zeolite-L and zeolite-Y. Obviously these are consequences of the different framework structure of the host; zeolite-Y cavities are large enough to accommodate clusters of Se¹³. Such prospects could well lead to the generation of novel electronic and optical behaviour.

O.T. thanks the Ministry of Education, Science and Culture for financial support and the British Council for an Academic

Travel Grant. O.T. and D.W. thank the Japan Securities Scholarship Foundation for financial aid.

Received 11 May; accepted 28 September 1987.

1. Bogomolov, V. N. *et al. JETP Lett.* **31**, 378-381 (1980).
2. Rabo, J. A. & Kasai, P. H. *Prog. Solid St. Chem.* **9**, 1-20 (1975).
3. Barrer, R. M. *Zeolites and Clay Minerals* (Academic, London, 1978).
4. Edwards, P. P. *et al. J. chem. Soc. chem. Commun.* 982-984 (1984).
5. Cotts, R. A., Carpenter, T. A., Edwards, P. P., Klinowski, J. & Thomas, J. M. *Zeolites Prog.* (submitted).
6. Bogomolov, V. N., Khodlodkevich, S. V., Romanov, S. G. & Agroskin, L. S. *Solid St. Commun.* **47**, 181-182 (1983).
7. Tamura, K., Hosokawa, S., Endo, H., Yamasaki, S. & Oyanagi, H. *J. phys. Soc. Japan* **55**, 528-533 (1986).
8. Bogomolov, V. N., Poborchh, V. V., Kholodkevich, S. V. & Shagin, S. I. *JETP Lett.* **38**, 532-535 (1984).
9. Goodman, P. & Moodie, A. F. *Acta crystallogr.* **A30**, 280-290 (1974).
10. Jefferson, D. A., Millward, G. R. & Thomas, J. M. *Acta crystallogr.* **A32**, 825-828 (1976).
11. Teraski, O., Thomas, J. M. & Millward, G. R. *Proc. R. Soc. A* **305**, 153-182 (1984).
12. Meier, W. M. *Z. Kristallogr.* **115**, 439-446 (1961).
13. Parise, J. B. *8th Australian National Chem. Conv.* Sydney, 25 August (1987). Presented orally at Univ. of New South Wales.

Dwarfism in an adolescent from the Italian late Upper Palaeolithic

David W. Frayer*, William A. Horton†, Roberto Macchiarelli‡ & Margherita Mussi§

* Department of Anthropology, University of Kansas, Lawrence, Kansas 66045, USA

† Medical Genetics Program, Department of Pediatrics, University of Texas, Houston, Texas 77030, USA

‡ Sezione di Antropologia, Museo Preistorico-Etnografico "L. Pigorini", Viale Le Lincoln 1, 00144 Rome, Italy

§ Dipartimento di Scienze dell'Antichità, Università di Roma "La Sapienza", Via Palestro 63, 00185 Rome, Italy

There have been numerous reports of pathological conditions in the hominid fossils, but these have only involved trauma or age-related deterioration in the health of otherwise normal individuals¹⁻⁴. Here we describe a skeleton of a young male from Riparo del Romito in Calabria, dated to the Epi-Gravettian of southern Italy. The preserved skeletal elements show that this individual (Romito 2) had the skull and long-bone morphology consistent with a mesomelic form of dwarfism, most probably the autosomal recessive disorder acromesomelic dysplasia⁵⁻⁸. Generally recognized at birth, persons with acromesomelic dysplasia usually have normal intelligence and are free of serious medical problems. However, growth deficiency is severe (adult height typically is 110-120 cm) and mobility at the elbows is restricted. These physical impairments would have greatly interfered with the individual's participation in subsistence activities and would have been a substantial handicap in a nomadic hunting and gathering group. Thus, besides being the earliest known case of dwarfism in the human record, this skeleton provides evidence of tolerance of, and care for, a severely deformed individual in the Palaeolithic.

During excavations between 1963 and 1965 at Riparo del Romito near Papisidero in Southern Italy, Graziosi discovered reasonably complete skeletons of six individuals in four graves from different areas of the cave^{9,10}. Two skeletons were found in a double burial below undisturbed Epi-Gravettian levels^{9,11}. Association of two large horn fragments with the skeletons and radiocarbon dates of $11,150 \pm 150$ years before the present (yr BP) for the grave are consistent with the remains being from the late Upper Palaeolithic^{9,12}. The two individuals represent an old female (Romito 1) and a late-adolescent male (Romito 2). The female is small in stature, but shows no skeletal abnormalities, whereas cranial and postcranial elements of the male are clearly pathological (Fig. 1).

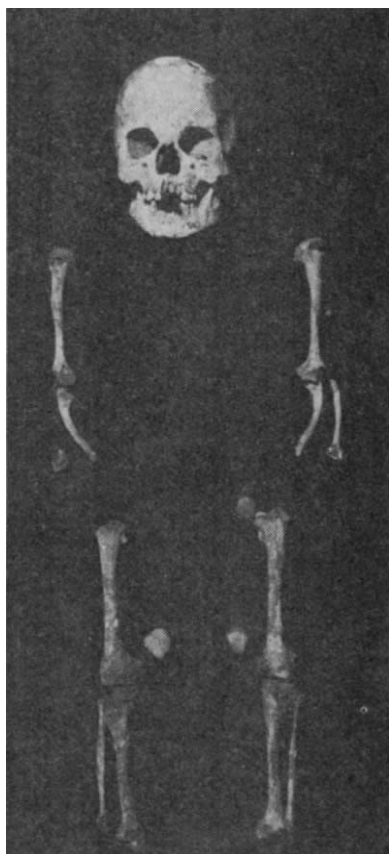


Fig. 1 The skeletal remains of Romito 2 (~10% natural size).

We base our diagnosis on specific skeletal indicators typical of acromesomelic dysplasia. These include severely reduced long-bone diaphyseal lengths resulting in a disproportionately short stature, extreme shortening and bowing of the radius and ulna with restricted elbow extension, a large rounded cranium with marked frontal and parietal bossing, a constricted cranial base and a depressed nasal root⁵⁻⁸. For Romito 2, all long bone epiphyses are fused, but maximum lengths are greatly reduced compared to individuals of normal stature from the Italian late Upper Palaeolithic and early Mesolithic (Table 1). In the upper extremity the forearm bones are very short, whereas in the lower limbs, the femur is reduced relative to the tibia. From the lengths of the tibia and femur, we estimate a stature of between 1.0 and 1.3 m, which is more than four standard deviations below the means for either the Upper Palaeolithic or Mesolithic in all of Europe^{13,14}. In addition to size, articulation of the complete left humerus and ulna show the individual could not have extended his elbow fully, but rather was capable of only a 130° extension.

As to the cranial features, the transverse dimensions of the skull base and lengths from the foramen magnum to the face are reduced, presumably related to retarded growth of the spheno-basilar region. The frontal bone is high, arching anteriorly over the orbits because of greatly expanded frontal bosses. The parietals also show substantial bossing, so that the greatest breadth of the skull occurs high on the vault. The cranial index (maximum cranial breadth divided by length) for this specimen of 82.5 exceeds published values for all other Upper Palaeolithic or Mesolithic crania^{14,15}, indicating substantial brachycephaly. Finally, the superior nasal region is moderately depressed. Thus, cranial and postcranial features of Romito 2 closely accord with conditions typical of acromesomelic dysplasia. The hands and feet, which are distinctive in this disorder, were unfortunately not preserved. However, no other known syndrome fits all the pathological features found in

Table 1 Long-bone dimensions of Romito 2 compared to Italian Upper Palaeolithic and Mesolithic specimens of normal stature

Sample	Romito 2	Comparative sample*		Romito 2 comparative sample (%)
		Mean	Range	
Maximum lengths				
Humerus	171	297	262-325	58
Radius	(102)	228	195-250	45
Ulna	111	247	222-268	45
Femur	225	426	370-473	53
Tibia	211	367	355-385	58
Indices†				
Radius humerus	60	77	74-81	
Ulna/humerus	65	83	79-85	
Tibia/femur	94	84	81-88	
Fibula/Femur	81	76	—	

All measurements are in mm.

* Includes males and females from Arene Candide, Barma Grande, Grotte des Enfants, Paglicci, Romito and Uzzo.

† Expressed as percentages.

Romito 2, so that acromesomelic dysplasia is the most likely diagnosis.

Although other cases of dwarfism have been described in European populations, these all occur in large cemeteries dated to the Middle Ages¹⁶⁻¹⁸. Other Old World and North American cases also derive from graveyards associated with agricultural or complex societies^{19,20}. Consequently, this individual provides the first example of dwarfism in a prehistoric hunter-gatherer sample and, moreover, extends the known antiquity of this type of inherited bone disease by more than 5,000 years. Acromesomelic dwarfism is very rare in living groups, so its occurrence in the Upper Palaeolithic is remarkable considering the low population densities estimated for these groups²¹. In addition, that the syndrome is inherited as an autosomal recessive⁷ suggests some degree of inbreeding during this period.

Finally, unlike other pathological conditions found in fossil hominids, where initially healthy individuals suffered diseases resulting from trauma or ageing, Romito 2 was profoundly different throughout his life. Sometime after birth, but certainly by late childhood, the individual must have been recognized as being of significantly different stature from his peers. Despite this, he must have been supported by members of his social group. Although individuals of severely limited stature lead relatively normal lives in modern society, they are constantly confronted with a physical world which is designed for average-size people²². This problem must have been a serious handicap in the Palaeolithic, given the demands of subsistence in a rigorous environment and of a nomadic life. For example, the reduced stature and lack of complete forearm extension would have greatly limited this individual's ability to participate effectively in hunting. Also, Mørch²³ reports that individuals of limited stature often have impaired mobility and that many tire after walking even short distances. This impediment would have further decreased the likelihood of effective hunting and even impinged on the ability of the individual to keep up with the group's periodic movements, especially in the rugged physical environment of southern Italy. Yet, despite these handicaps, the dwarf survived to about 17 years of age. Moreover, once he died, he was accorded special funereal treatment as evidenced by his burial in an important cave. Riparo del Romito has a long prehistoric occupation, occupies a strategic location in a valley linking the Tyrrhenian Sea to the Ionian Sea, and is one of the few caves in Italy with parietal art^{24,25}. Thus, the site was an important social and/or ritual centre and, it is clear from the limited number of burials in the cave that only certain members of the population were of the status necessary to be

buried there. Romito 2 received such treatment which may also attest to his acceptance by the group despite his severe handicaps and limited ability to contribute to subsistence and other economic activities.

We thank the Soprintendenza Archeologica della Calabria, Dr Elena Lattanzi, for permission to study the Romito 1 and 2 skeletons at the Museo Nazionale di Reggio Calabria and Dr Emilia Andronico (Reggio Calabria), I. Pap (Budapest) and E. Iregren (Lund) for their help. D.W.F. was supported by an NSF grant.

Received 9 September; accepted 24 September 1987.

- DeLumley, M.-A. *Études Quat.* **2**, 1-626 (1973).
- Dastague, J. & deLumley, M.-A. in *La Préhistoire Française* (ed. deLumley, H.) 612-622 (CNRS, Paris, 1976).
- Corrain, C., Grazzi, G. & Leonardi, P. *Preist. Alpina* **12**, 175-212 (1976).
- Pittard, E. & Sauter, M.-R. *Archs suisses Anthropol. gén.* **XI**(2), 1-197 (1946).
- Maroteaux, P., Martinelli, B. & Campailla, E. *La Presse Méd.* **79**, 1839-1842 (1971).
- Langer, L. O. *et al. Am. J. med. Genet.* **1**, 87-100 (1977).
- Langer, L. O. & Garrett, R. O. *Radiology* **137**, 349-355 (1980).
- Rimoin, D. L. *Adv. hum. Genet.* **5**, 1-118 (1975).
- Graziosi, P. *Riv. Sci. preist.* **18**, 315 (1963).
- Graziosi, P. *Riv. Sci. preist.* **18**, 301-302 (1964).
- Messeri, P. *Atti X Riun. scient. Ist. Ital. Preist. Protost.* 301-307 (1966).
- Alessio, M., Bella, F. & Cortesi, C. *Radiocarbon* **8**, 401-412 (1966).
- Frayer, D. W. *Am. Anthropol.* **83**, 53-73 (1981).
- Vallois, H. V. & deFélice, S. *Arch. l'Inst. Paléont. hum.* **37**, 1-193 (1937).
- Riquet, R. in *L'Homme, de Crô-Magnon* (eds Camps, G. & Olivier, G.) 37-58 (Arts et Métiers Graphiques, Paris, 1970).
- Farkas, G. & Lengyel, I. *Móra Ferenc Múzeum Évkönyve*, 199-207 (1971/1972).
- Gładkowska-Rzeczycka, J. *Ossa* **7**, 71-74 (1980).
- Larje, R. *Arch. Environ.* **4**, 259-271 (1985).
- Johnston, F. E. *Clin. Pediat.* **2**, 703-708 (1963).
- Ortner, D. J. & Putschar, W. J. *Smithson. Contr. Anthropol.* **28** 1-479 (1981).
- Wobst, H. M. in *The Demographic Evolution of Human Populations* (eds Ward, R. H. & Weiss, K. M.) 49-58 (Academic, New York, 1976).
- Ablon, J. *Little People in America* (Praeger, New York, 1984).
- Mørch, E. T. *Chondrodystrophic Dwarfs in Denmark* (Munksgaard, Copenhagen, 1941).
- Graziosi, P. in *Miscelánea en Homenaje al Abate Henri Breuil* (ed. Perello, E. R.) 457-466 (Deputación Provincial, Barcelona, 1964).
- Graziosi, P. *L'Arte Preistorica in Italia* (Sansoni, Firenze, 1973).

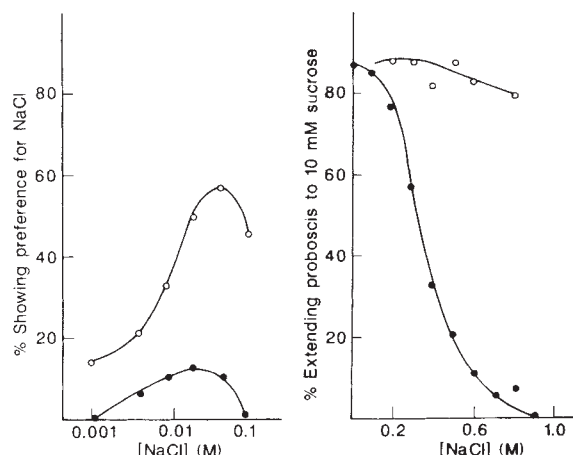


Fig. 1 Behavioural responses of wild-type and *gustB* flies to NaCl. *a*, Feeding preference test: The wells of a 6 × 10 microtitre plate were filled with 1% agar solution. Alternate wells also contained 0.2% of the food dye carmoisine red; uncoloured wells contained increasing concentrations of NaCl. Freely moving flies were placed on these plates for 90 min after which the flies with uncoloured abdomens were counted. The percentage of uncoloured flies measures preference for NaCl: 18% of the flies remain uncoloured in the absence of salt. These are presumably non-feeders. This value was subtracted from measurements. There is no preference for the dye. *b*, Inhibitory effect of NaCl on proboscis extension². Flies were starved in a moist chamber for 18 h and water-satiated before testing. The tarsi of prothoracic legs were touched with a solution of 10 mM sucrose containing increasing amounts of NaCl. Each fly was tested 10 times and 7 or more extensions of proboscis were counted as an acceptance response. ○, *Gust B*, ●, wild type.

A gene affecting the specificity of the chemosensory neurons of *Drosophila*

Kavita Arora, Veronica Rodrigues, Swati Joshi, Shubha Shanbhag & Obaid Siddiqi

Molecular Biology Unit, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400 005, India

The sensilla on the proboscis and tarsi of *Drosophila* contain five neurons, four chemosensory and one mechanosensory¹⁻³. The sugar-sensitive neuron, designated S, carries independent acceptor sites for pyranose, furanose and trehalose⁴⁻⁷. Two others, L1 and L2, respond to salts². The fourth neuron, W, is inhibited by salts and sugars, and is believed to mediate detection of water. We describe here a gene in which mutations alter the neurons in such a way that the S cell is excited by salts. As a result, the mutant flies are strongly attracted by NaCl at concentrations which are repellent to the wild type. To our knowledge, this is the first instance of a mutation which changes the specificity of the chemosensory neurons.

Wild-type flies show a slight preference for concentrations of NaCl up to 10 mM. Salt concentrations above this inhibit ingestion. The preference of *gustB* mutant flies for NaCl in such feeding tests is markedly greater (Fig. 1A). A similar preference for NaCl is evident in proboscis extension tests. The proboscis extension response evoked by 10 mM sucrose in *gustB* shows an increased tolerance to NaCl (Fig. 1B).

The gene *gustB* is located at 1.38 between *vermillion* and *forked* on the X chromosome. Two ethylmethane-sulphonate-induced alleles, ×5 and ×7 were isolated in our laboratory; a third allele was provided by Dr Laurie Tompkins. Mutant *gustB* alleles are recessive and fall into a single complementation group.

Heterozygotes of *gustB* × 5 with *Df(1)m259-4* (polytene chromosome bands 10C12-10E3) and *Df(1)KA6* (10E1-11A7) have the *gustB* phenotype, placing the gene in the bands 10E1, 2. The three alleles and their heterozygotes with *Df(1)KA6* have similar phenotypes.

In the wild-type Canton *S* strain of *Drosophila*, each half labellum carries 34 gustatory hairs^{1,2} arranged in about four rows. There are two medial rows, one of short and one of long (*M*-type) sensilla, a central irregular row of long and intermediate length (*I*-type) sensilla and a peripheral row of six short (*P*-type) sensilla⁸. Typically, the hairs of the medial and central rows contain all four chemosensory neurons S, L1, L2 and W. Most *P*-type hairs and some of the *I*-type hairs contain only two chemosensory neurons, responding to sugars and salts respectively.

We examined the electrophysiological responses of the label hairs of *gustB*, using the tip recording method of Hodgson *et al.*⁹. The spikes from the four chemosensory neurons can be distinguished by their relative amplitudes²: *gustB* mutants show enhanced firing when stimulated by NaCl (Fig. 2A). Examination of traces from *M*-hairs of the medial row of *gustB* mutants showed that the spikes evoked by NaCl were markedly heterogeneous; a group of spikes were present that were somewhat smaller than the L1 spikes and seemed to resemble the S spikes (Fig. 2b).

Spike amplitudes vary from hair to hair and depend on the strength of the electrolyte or the capacitance of the recording probe. More importantly, amplitudes increase with the frequency of firing^{8,10}. The absolute height of the spike is, therefore, not a reliable guide to the identity of a neuron. On the other hand, the relative amplitudes of spikes from the neurons of a sensillum vary much less, and when conditions of stimulation are appropriately chosen, the neurons can be identified with confidence. L2 and W are easy to recognize but the distinction between S and L1 requires careful analysis.

We made recordings from 63 *M*-type hairs of 20 *gustB* females

Fig. 2 Extracellular sensory responses of wild-type and *gustB* flies. The labellar hairs were touched with the stimulating solution in a capillary which also served as the recording electrode. Impulses were counted for a period of 1 s starting 50 ms after contact. **A**, Response of an M-type hair to NaCl, ●, normal flies; ○, *gustB* × 5. All spikes were counted. **B**, L and S spikes from wild-type (CS) and *gustB* flies. Top to bottom: a, CS, 4 mM NaCl (L1 spikes); b, CS, 4 mM sucrose in 0.5 mM NaCl (S spikes); c, CS, 2 mM NaCl + 2 mM sucrose (L and S spikes); d, *gustB*, 2 mM NaCl (L and S spikes); e, *gustB*, 4 mM sucrose in 0.5 mM NaCl.

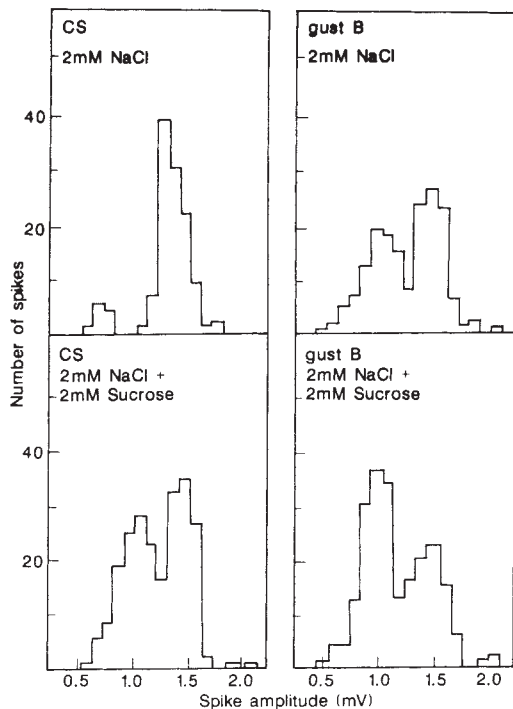
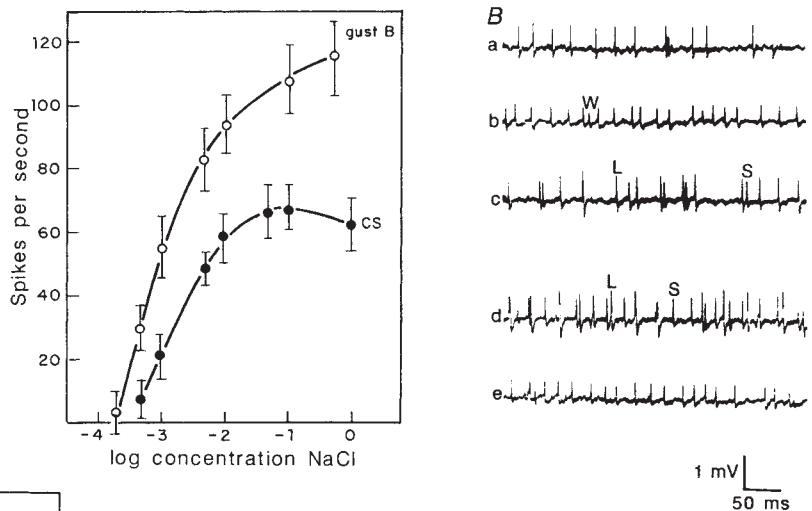


Fig. 3 Amplitudes of NaCl-induced spikes in wild-type and *gustB* flies: histograms show the amplitude distribution of spikes from a single sensillum in wild-type (left panels) and *gustB* (right panels). In these hairs the L and S spikes were clearly distinguishable. NaCl alone (upper panels) excites the S in addition to the L cell in *gustB* (upper right); addition of sucrose (lower panels) stimulates the S cell in both Canton S and *gustB*.

and compared these with Canton S controls. In 22 of these *gustB* hairs the L1 spikes were greater than the S spikes by more than 0.4 mV and the two types of spikes could be classified by inspection. The amplitude distribution was clearly bimodal (Fig. 3). In addition to exciting L1, NaCl evokes a spike which has the same amplitude as the sugar-stimulated spike. In the remaining 41 hairs the unusual heterogeneity of the NaCl-induced spikes was evident from the distribution of their amplitudes. The L1 and S spikes in the wild type were distributed unimodally with a standard deviation of 8.3%; *gustB* stimulated with 10 mM NaCl yielded a broad distribution shifted and skewed towards the S spike with a standard deviation of 13.3% (data not shown). The expression of the *gustB* phenotype is thus not restricted to a selected sub-population of hairs.

The mutation affects sensilla of all types, but its effect is most clearly demonstrable in the P-type hairs. The L neuron in the P-type hairs produces a spike of about 1.5 mV with a concentration threshold for NaCl above 100 mM. The firing of the S

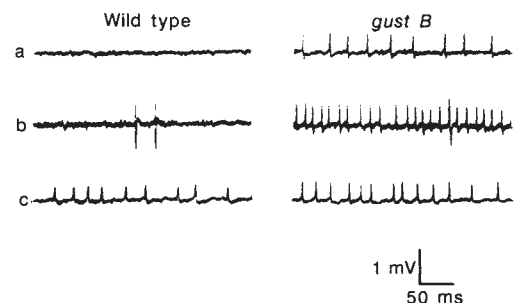


Fig. 4 Response of the S neuron in P-type hairs to NaCl and sucrose. Left, wild type; right *gustB*. a, 2 mM NaCl; b, 100 mM NaCl; c, 10 mM sucrose. The larger spikes in the b traces are from the single L neuron of the P-type hair.

neuron, stimulated by 10 mM NaCl can, therefore, be seen without the background of L1 spikes present in the M-type hairs (Fig. 4). The spikes induced by 10 mM NaCl in the P-type hairs of *gustB* are quite different from the L spikes and indistinguishable from the S spikes.

It might be argued that the unexpected spikes come from an additional neuron not present in the normal flies. We examined the *gustB* labellar hairs by transmission electron microscopy. A sample of 137 hairs of *gustB* showed no departure from the neuronal composition of the wild type⁸. The branching pattern and the dendritic morphology of neurons of the mutant were not recognizably altered.

These results are part of our efforts to dissect genetically the gustatory receptor system of *Drosophila*. Mutants with altered gustatory responses have been reported by several groups^{2,5,11,12}. So far, *gustB* is the only instance of a gene whose mutations enhance the sensitivity of a particular gustatory neuron and change its specificity. The gene seems to be involved in a regulatory function concerned with the segregation of cationic sensitivity in gustatory neurons. The fact that all the known *gustB* alleles are recessive implies that this regulation is negative.

Received 11 June; accepted 14 September 1987.

- Falk, R., Bleiser-Avivi, N. & Atidia, J. *J. Morph.* **150**, 327-341 (1976).
- Rodrigues, V. & Siddiqi, O. *Proc. Indian Acad. Sci. B* **87**, 147-160 (1978).
- Nayak, S. V. & Singh, R. N. *Int. J. Insect. Morph. Embryol.* **12**, 273-291 (1983).
- Tanimura, T. & Shimada, I. *J. comp. Physiol.* **141**, 265-269 (1981).
- Tanimura, T., Isono, K. & Takamura, I. *J. comp. Physiol.* **147**, 433-437 (1982).
- Rodrigues, V. & Siddiqi, O. *Molec. gen. Genet.* **181**, 406-408 (1981).
- Dethier, V. G. *The Hungry Fly* (Harvard University Press, Massachusetts, 1976).
- Arora, K. thesis, Bombay Univ. (1985).
- Hodgson, E. S., Lettvin, J. Y. & Roeder, K. D. *Science* **122**, 417-418 (1955).
- Fujishiro, N., Kijima, H. & Morita, H. *J. Insect Physiol.* **30**, 317-325 (1984).
- Falk, R. & Atidia, J. *Nature* **254**, 325-326 (1975).
- Tompkins, L., Cardosa, J., White, F. & Sanders, T. G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 884-887 (1979).

Co-release of glutamate and aspartate from cholinergic and GABAergic synaptosomes

Maureen Docherty, H. F. Bradford & J.-Y. Wu*

Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, UK

* Department of Physiology, Pennsylvania State University, College of Medicine, Hershey, Pennsylvania 17033, USA

Glutamate and aspartate are known to be released in a calcium-dependent fashion by depolarizing stimulation of mammalian brain synaptosomes (isolated nerve endings)¹⁻³, an observation which strengthens their claims to be neurotransmitter candidates. The source of these compounds has been interpreted as the exclusively glutamatergic or aspartatergic synaptosome sub-populations assumed to be present in the standard heterogeneous preparations from mammalian brain. Several neurotransmitter-specific synaptosomal surface markers have recently been identified by immunolysis studies⁴⁻⁷ and these have allowed separation of subpopulations of synaptosomes by an affinity purification method⁸. These markers appear to be closely related to the biosynthetic enzyme for the principal neurotransmitter released by each sub-category of synaptosome. We have isolated highly purified, metabolically active, GABAergic and cholinergic synaptosomes from cerebral cortex using antisera recognizing either glutamate decarboxylase (GAD) or choline acetyltransferase (ChAT), in conjunction with magnetic microspheres covalently coupled to Protein A (ref. 8), and now report that these synaptosomes release both glutamate and aspartate, in addition to their principal neurotransmitter, when treated with chemical depolarizing agents.

Whole synaptosome fractions from cerebral cortex were incubated with non-immune (as control) or immune serum to allow specific immunoglobulins to attach to the synaptosomal surface as appropriate. Labelled synaptosomes were then exposed to magnetic microspheres to which protein A had been conjugated. Synaptosomes could then be collected with the aid of a magnet⁸. Using this immunomagnetophoretic technique it is possible to obtain highly purified, metabolically active, cholinergic and GABAergic synaptosomes from rat cerebral cortex⁸ (Table 1). Following multiple collection of anti-ChAT labelled synaptosomes, 52% of the total ChAT and 15% of ACh were recovered in a fraction containing about 5% of the total lactate dehydrogenase (LDH), which was devoid of GAD activity (Tables 1 and 2). When anti-GAD labelled synaptosomes were purified, 40% of the total GAD and 22% of γ -amino-butyric acid (GABA) was recovered along with 15% of total LDH (Tables 1 and 2). This GABAergic fraction showed a slight contamination with ChAT (1.8% of total present in mother fraction, Table 1). When compared with the mother fraction, the cholinergic subfraction was found to contain both elevated aspartate and glutamate levels (about fourfold higher) compared with the 'metabolic' amino acids threonine or serine (Table 2). The GABAergic subfraction also showed raised aspartate (2-3-fold higher) and glutamate levels (6-9-fold higher on this basis). When these figures were expressed in terms of the LDH content of each subfraction (namely as an index of the number of synaptosomes present, Table 3) and compared to this ratio in the mother fraction, it could be seen that the cholinergic synaptosomes were enriched in acetylcholine (ACh), choline, aspartate and glutamate, and that the GABAergic synaptosomes were enriched in both glutamate and GABA, but not in aspartate. ChAT was enriched tenfold and GAD threefold on this basis in cholinergic and GABAergic subpopulations respectively, which is close to the theoretical maximum for each^{4,8}.

Some of these constituents could be released from the incubated synaptosomes by stimulation with either potassium

Table 1 Lactate dehydrogenase, choline acetyltransferase and glutamate decarboxylase activity in cholinergic and GABAergic synaptosomes

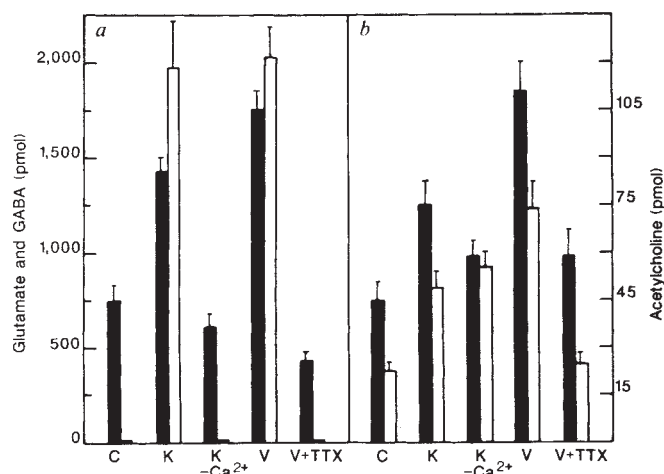
Content per fraction	Whole preparation	Cholinergic synaptosomes	GABAergic synaptosomes
LDH nmol min ⁻¹	2,426 $\pm$ 301	130 $\pm$ 24	372 $\pm$ 43
ChAT pmol min ⁻¹	4,096 $\pm$ 540	2,127 $\pm$ 270	75 $\pm$ 10
GAD pmol min ⁻¹	1,532 $\pm$ 201	ND	620 $\pm$ 80

Synaptosomes prepared from rat cerebral cortex. Each value represents the mean $\pm$ s.d. for quadruplicate batch separations using synaptosomes from three separate synaptosome preparations. Non-specific binding, estimated using non-immune serum (1:20), amounted to 3-4% of the specific binding. ND, not detectable. Cholinergic or GABAergic synaptosomes⁸ from rat cerebral cortex⁹ were prepared by specific cell-surface labelling with anti-ChAT (1:20) or anti-GAD (1:40) serum⁴ respectively (at 0-5 °C for 45 min in Krebs-phosphate buffer containing glucose, 30 mM (KPBG). After washing by sedimentation (10,000g for 30 s) and resuspension in KPBG or 0.32M glucose containing bovine serum albumin (0.01% w/v), each batch of cortical synaptosomes (3 mg protein) was incubated using a roller mixer for 45 min at 0-5 °C with 125 μ l Magnogel (Act-Magnogel AcA 44, LKB) which had been conjugated to Protein A (2 mg ml⁻¹ of Magnogel). The Magnogel was then rapidly (4-5 s) sedimented (using a horse-shoe magnet) to produce supernatant S1. The Magnogel was washed by resuspension and resedimentation in the magnetic field. Fresh Magnogel-Protein A (125 μ l) was then added to the supernatant S1 and incubated for a further 45 min at 0-5 °C. After sedimentation in the magnetic field (yielding a second pellet) synaptosomes remaining in the supernatant (S2) were exposed for a third time to fresh Magnogel-Protein A. All pellets were washed by resuspension in KPBG or 0.32M glucose and resedimentation in the magnetic field before assay. Synaptosomes were removed from the affinity gel by addition of Triton X-100 (1%, w/v) and LDH and ChAT were assayed essentially as described before⁴. GAD activity was assayed⁴ after hypotonic lysis with 50 mM sodium phosphate buffer, pH 7.4, containing pyridoxal phosphate (0.2 mM) and freezing rapidly and thawing (twice).

(50 mM) or veratrine (75 μ M) (Fig. 1). Thus, in response to these depolarizing treatments, GABAergic synaptosomes released glutamate (Fig. 1) and aspartate (potassium stimulation, 40% increase; with veratrine, 130% increase above control) in addition to GABA, and cholinergic synaptosomes released these two amino acids (aspartate release, 300% increase with potassium; 400% increase with veratrine above control) as well as ACh, though neither subpopulation showed enhanced release of taurine. The release of glutamate and aspartate (though not GABA in these experiments) stimulated by high potassium appeared to be calcium-dependent but that evoked by veratrine was calcium-independent, as reported by other workers¹¹. Tetrodotoxin (TTX) treatment in most cases abolished veratrine-stimulated release, but, as expected, it was ineffective against that evoked by high potassium. These findings indicate that the stimulus-induced efflux of these compounds possessed properties typical of those found for conventional synaptosome preparations and brain slices, and usually taken as evidence of its physiological significance^{11,12}.

As the precursor of GABA, glutamate enrichment in GABAergic synaptosomes could be expected, but on an LDH basis it is more enriched in the cholinergic type. Its stimulus-evoked release from these synaptosomes suggests it could act as a neurotransmitter or neuromodulator in both categories of nerve terminal.

It is conceivable that the observed glutamate and aspartate release originates from separate (for example, glutamatergic or aspartatergic) subpopulations which are accidentally co-purified. The fact that 86% of LDH and 42% of the ChAT in cholinergic synaptosomes attached to the Magnogel could be released by immunolysis using active complement (Table 3), indicates that most of the synaptosomes carried membrane-bound ChAT and were therefore likely to be cholinergic in



nature (compare with 10–15% LDH and about 40% ChAT released from the mother fraction⁴). Also, there does not seem to be any higher degree of nonspecific binding of aspartate or glutamate-containing terminals as less than 0.1% of these amino acids was found attached to the Magnogel-protein A microspheres.

These results indicate that glutamate can be released from cholinergic and GABAergic terminals. Is it possible that glutamate is released as a neurotransmitter or neuromodulator from a wide category of neurons, part at least being directly from the synaptoplasm^{12–14}. Exclusively glutamatergic neurons are possibly far fewer than implied by the widespread occurrence of glutamate receptors^{15,16} and the high sensitivity of most neurons to iontophoresed glutamate¹⁷. Electrophysiological findings would not conflict with this proposal¹⁸. In line with such co-release of amino acids and 'classical' neurotransmitters is the recent report that pure cholinergic synaptosomes from *Torpedo* release glutamate along with ACh (ref. 19).

Even the apparent anomaly that an excitatory (glutamate) and an inhibitory (GABA) neurotransmitter are released from

Fig. 1 Stimulated release of glutamate, GABA, and ACh from *a*, cholinergic and *b*, GABAergic synaptosomes, prepared as described in Table 2 legend. Values represent the mean $\pm$ s.d. of responses for 5–7 batch separations from four separate synaptosome preparations. C, Control; K, 50 mM KCl added; K–Ca²⁺, 50 mM KCl, no Ca²⁺ present; V, veratrine (75 μ M); V+TTX, veratrine (75 μ M) plus tetrodotoxin (1 μ M). Solid bars represent glutamate release; open bars *a*, ACh; *b*, GABA release.

Methods. Cholinergic and GABAergic synaptosomes (from 3 mg protein, cortical synaptosomes) still attached to Magnogel-Protein A, were resuspended in 500 μ l of either KPBG or KPBG medium without calcium (KPBG containing no CaCl₂, MgSO₄ at 10 mM and EGTA at 20 mM) and preincubated at 37°C for 15 min with manual mixing about every 5 min. Synaptosome suspensions were then stimulated for 15 min by the addition of the appropriate KPBG containing either elevated potassium (50 mM) or veratrine (75 μ M), addition of which increased the incubation volume to 1 ml. In some experiments synaptosomes were exposed to tetrodotoxin (TTX, 1 μ M) 5 min before stimulation. At the end of the incubation period, synaptosomes attached to Magnogel were sedimented in a magnetic field, supernatants were removed, and pelleted synaptosomes were washed in the appropriate KPBG (1 ml). Amino-acid analysis was performed on both supernatants and pellets as described in the Table 2 legend.

the same terminal could be encompassed by a mechanism such as one allowing differential release of two neuroactive substances at different stimulus frequencies²⁰, or the modulation of the release of one neurotransmitter by the other, through activation of an appropriate presynaptic receptor linked directly to the transmitter release process²¹.

In addition, our own observations on complement-mediated immunolysis of whole cortical synaptosome preparations have shown that at least 66% of synaptosomes could be lysed, as judged by summation of maximal LDH release resulting from immunolysis mediated with anti-ChAT, anti-GAD and anti-dopamine β -hydroxylase^{4,6,7}. This would leave a residue of only about 34% of synaptosomes which would have to accommodate the supposedly large glutamatergic subpopulation, together with the peptidergic, serotonergic, histaminergic, adrenergic, purinergic, and other subtypes known to be present in the cerebral cortex, assuming that all these neurotransmitters are not co-released. A similar high degree of lysis (69%) of striatal synaptosomes can be achieved using antisera to ChAT, GAD, tyrosine hydroxylase and tryptophan hydroxylase^{4–7}, and therefore

Table 2 Amino-acid, acetylcholine (ACh) and choline content of purified cholinergic and GABAergic synaptosomes

Content per fraction (pmol/separation)	Total mother fraction	Cholinergic subpopulation	% Total	GABAergic subpopulation	% Total
Taurine	28,552 $\pm$ 3,061	731 $\pm$ 60	2.5	696 $\pm$ 80	2
Aspartate	26,456 $\pm$ 2,931	1,986 $\pm$ 221	8	1,631 $\pm$ 179	6
Threonine	5,459 $\pm$ 632	120 $\pm$ 15	2	142 $\pm$ 23	3
Serine	8,321 $\pm$ 769	191 $\pm$ 21	2	187 $\pm$ 30	2
Glutamate	41,994 $\pm$ 4,063	3,359 $\pm$ 394	8	7,559 $\pm$ 893	18
Glutamine	7,139 $\pm$ 921	0	—	0	—
Glycine	8,819 $\pm$ 763	96 $\pm$ 12	1	75 $\pm$ 12	1
Alanine	7,063 $\pm$ 839	106 $\pm$ 20	1.5	103 $\pm$ 16	1
Valine	2,870 $\pm$ 321	72 $\pm$ 15	2.5	93 $\pm$ 12	3
Isoleucine	1,210 $\pm$ 213	low	—	low	—
Leucine	1,096 $\pm$ 313	low	—	low	—
Tyrosine	809 $\pm$ 110	low	—	low	—
Phenylalanine	605 $\pm$ 109	low	—	low	—
GABA	14,531 $\pm$ 1,658	320 $\pm$ 43	2	3,150 $\pm$ 429	22
Histidine	7,831 $\pm$ 902	121 $\pm$ 17	1.5	262 $\pm$ 35	3
ACh	1,380 $\pm$ 209	206 $\pm$ 31	15	41 $\pm$ 8	3
Choline	4,200 $\pm$ 563	621 $\pm$ 109	15	504 $\pm$ 83	12

Values represent means $\pm$ s.d. for triplicate batch separations from three separate synaptosome preparations. Values expressed as 'low' are <20 pmol per separation; —, not detectable.

Methods. Cholinergic and GABAergic synaptosomes were prepared as described in the Table 1 legend. Amino-acid analysis was performed on the first separated pellet (after addition of methanol (1 ml) containing an appropriate quantity of norleucine in 0.025N HCl as internal standard, and freezing to -20°C and thawing) as previously described⁴. No amino acids were extracted from Magnogel-Protein A plus antiserum, in the absence of synaptosomes using this procedure. Acetylcholine was measured as previously described¹⁰.

Table 3 Purification of amino acids, ACh, choline, ChAT and GAD with respect to LDH content in cholinergic and GABAergic synaptosomes

a	Substance	Purification factor	
		Cholinergic synaptosomes	GABAergic synaptosomes
	Aspartate	×2.67	×0.54
	Glutamate	×2.67	×1.64
	GABA	×0.67	×2.00
	ACh	×5.00	×0.27
	Choline	×5.00	×1.09
	Threonine	×0.67	×0.27
	Serine	×0.67	×0.18
	Alanine	×0.50	×0.09
	ChAT	×9.73	×0.11
	GAD	—	×2.63
b			
% Total released by immunolysis	ChAT	LDH	
Normal complement	42	86%	
Heat-inactivated complement	none	14%	

Ratios are calculated from % amino acid, ACh or choline *a*, and % LDH present in the synaptosome subpopulation, relative to that in the mother fraction. Thus, all ratios in the mother fraction are assigned as 1 (100%/100%); >1, increase in content, and <1, decrease in content relative to mother fraction. *b*, Complement-mediated immunolysis of the cholinergic synaptosomes was performed as described previously except that anti-ChAT was not added because the synaptosomes had already been pretreated with this antibody for separation purposes^{4,8}. ChAT and LDH release data are mean from immunolysis of three (normal complement) or two (heat-inactivated complement) batch separations.

similar arguments apply to this brain region. Such calculations would put the contribution of exclusively glutamatergic terminals at a relatively lower figure than would be predicted by the widespread occurrence of glutamate receptors and putative glutamatergic pathways^{15,16,18}. This, too, would therefore argue in favour of a function for glutamate as a cotransmitter or neuromodulator, and thus for a less frequent occurrence in the CNS of exclusively glutamatergic neurones compared with the current estimate¹⁶. Of course, evaluation of the extent to which these *in vitro* findings with synaptosomes reflect the properties of intact nerve-endings *in vivo* must await complementary *in vivo* studies.

The authors would like to thank Dr E. J. White and Professor J. B. Clarke for help with the assay of acetylcholine and choline.

Received 13 July; accepted 1 October 1987.

- Bradford, H. F. *Brain Res.* **19**, 239–247 (1970).
- Bradford, H. F. in *Chemical Neurobiology*, 229–232 (Freeman, New York, 1986).
- Bradford, H. F., Bennett, G. W. & Thomas, A. J. *J. Neurochem.* **21**, 495–505 (1973).
- Docherty, M., Bradford, H. F., Wu, J.-Y., Joh, T. H. & Reis, D. J. *Brain Res.* **339**, 105–113 (1985).
- Docherty, M., Bradford, H. F., Cash, C. D. & Maitre, M. *FEBS Lett.* **182**, 489–492 (1985).
- Docherty, M., Bradford, H. F. & Joh, T. H. *FEBS Lett.* **202**, 37–40 (1986).
- Docherty, M. & Bradford, H. F. *Neurosci. Lett.* **70**, 234–238 (1986).
- Docherty, M., Bradford, H. F. & Wu, J.-Y. *Neurosci. Lett.* **81**, 232–238 (1987).
- Dodd, P. R. *et al. Brain Res.* **226**, 275–287 (1981).
- Willoughby, J., Harvey, S. A. K. & Clarke, J. B. *Biochem. J.* **235**, 215–223 (1986).
- Norris, P. J., Dahlwal, D., Druce, D. & Bradford, H. F. *J. Neurochem.* **40**, 514–521 (1981).
- de Belleruche, J. S. & Bradford, H. F. *J. Neurochem.* **19**, 585–602 (1972).
- de Belleruche, J. S. & Bradford, H. F. *J. Neurochem.* **29**, 335–343 (1977).
- Nicholls, D. G. & Shira, T. S. *Nature* **321**, 772–773 (1986).
- Foster, A. C. & Fagg, G. E. *Brain Res. Rev.* **7**, 103–164 (1984).
- Monaghan, D. T. & Cotman, C. W. in *Excitatory Amino Acids* (eds Roberts, P. J., Storm-Mathisen, J. & Bradford, H. F.) 279–299 (Macmillan, London, 1986).
- Watkins, J. C. in *Kainic Acid as a Tool in Neurobiology* (eds McGeer, E. G., Olney, J. W. & McGeer, P. L.) 37–69 (Raven, New York, 1978).
- Meyer, M. L. & Westbrook, G. L. *Prog. Neurobiol.* **28**, 197–276 (1987).
- Vyas, S. & Bradford, H. F. *Neurosci. Lett.* **82**, 58–64 (1987).
- Lundberg, J. M., Anggard, A., Fahrenkrug, J., Lundgren, G. & Holmstedt, B. *Acta physiol. scand.* **115**, 525–528 (1982).
- Burnstock, G. in *Dale's principle and communication between neurones* (ed. Osborne, N. N.) 7–35 (Pergamon, Oxford, 1983).

Mediation of cell volume regulation by Ca^{2+} influx through stretch-activated channels

Ove Christensen

Department of General Physiology and Biophysics, The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

Animal cells initially swell in hypotonic media by osmotic water equilibration, but their volume is subsequently regulated by a net loss of KCl and amino acids with concomitant loss of cell water^{1,2}. Mechanisms for regulating cell volume are important in allowing cells to adapt to variations in external tonicity and metabolic load³. In red cells the KCl loss is mediated by electroneutral ion transport mechanisms⁴. In contrast, conductive K^+ and Cl^- transport pathways are activated during regulatory volume decrease in several cell types^{5,6} including epithelia^{7,8}. The activation seems to be mediated⁹ by internal Ca^{2+} , but the detailed mechanism is not known. In a leaky epithelium, the choroid plexus epithelium, we have found a cation-selective, Ca^{2+} -permeable channel which opens with membrane stretch^{10–12}. The epithelium also contains a high density of the large (~200 pS) type of Ca^{2+} - and voltage-activated K^+ channel¹³. Both channels are normally closed. I propose that in hypotonic media, the stretching of the cell membrane produced by the initial swelling causes influx of Ca^{2+} through the stretch-activated channels, which activates the neighbouring large K^+ channels to produce increased K^+ outflux with associated loss of cell water.

Choroid plexus epithelium was obtained from *Necturus Maculosus* an aquatic, gilled salamander from north America. The ionic channels in the ventricular membrane were investigated by the patch-clamp technique¹⁴, using the cell-attached

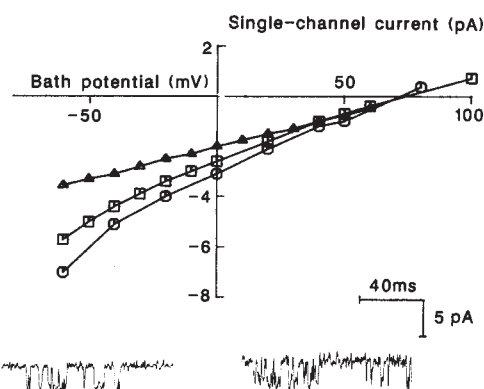


Fig. 1 Single-channel current versus bath potential recorded on cell-attached patches from the ventricular membrane of the choroid plexus epithelium. The electrodes contained 112 mM NaCl ($\square$), 112 mM KCl ($\circ$) or 75 mM CaCl_2 ($\triangle$). Negative currents indicate cation movement from electrode to cell. Insets below, typical time courses of single-channel current with K^+ (or Na^+) in the pipette and BP 0 mV (left) and with Ca^{2+} in the pipette and BP -40 mV (right) (1.2 kHz filtering).

Methods. The tissue was superfused with standard saline containing (mM): 104 NaCl, 8 NaHCO_3 , 2 KCl, 1 MgSO_4 and 1 CaCl_2 . The solution was bubbled by a mixture of 5% CO_2 and air, regulated to produce a pH of 7.4 at the tissue. Electrodes contained, apart from the stated concentrations of ions, 1 mM MgSO_4 and 10 mM HEPES, pH=7.4. Unless otherwise stated, the electrode concentrations were 112 mM. Cell potential was estimated from the reversal potential for the SAC or, with KCl in the electrode, for the large K^+ channel. For analysis, currents were filtered with the stated cut-off frequency (-3 dB; 24 dB per octave). The experiments were done at room temperature.

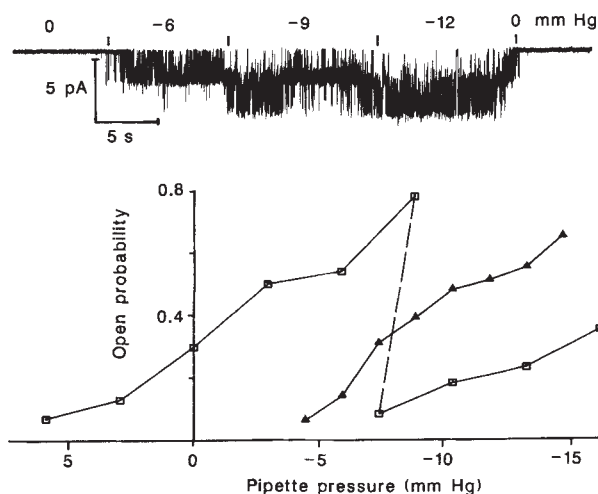


Fig. 2 Above, response of current through an attached patch to suction applied to the membrane through the electrode. (Electrode: KCl; 0 mV BP (=resting cell potential); 300 Hz filtering). Below, open probability for two other patches as function of pipette pressure (suction). The cell-attached data (□) were obtained with KCl in the electrode at 0 mV BP; dashed line, spontaneous shift of open probability. The data for the inside-out patch (Δ) were obtained with KCl in the electrode at -60 mV inside potential. Inside solution was 112 mM KCl, 1 mM MgSO_4 , 10 mM HEPES, pH 7.4. Open probability was calculated by computer analysis¹³ or by low-pass filtering (~1 Hz) of the current signal.

and the inside-out configurations. Potential across the patch was controlled by changing the bath potential (BP), the electrode being at zero potential¹³. For cell-attached patches, membrane potential (MP) equals bath potential plus the cell potential of -80 to -90 mV (ref. 15).

In cell-attached patches, the stretch-activated channel (SAC) was usually spontaneously active. Identical current-voltage (I-V) curves were obtained with 112 mM NaCl or KCl in the electrode (Fig. 1). There was a pronounced curvature of the I-V curves towards negative bath potentials. The conductance was 54 ± 3 pS at -20 to -60 mV, 40 ± 3 pS at 0 mV and 29 ± 2 pS at bath potentials larger than 40 mV (mean $\pm$ s.e.m., n between 10 and 14). With 75 mM CaCl_2 in the electrode, the conductance was 20 ± 1.6 pS ($n = 5$). In inside-out patches, the channel was rarely spontaneously active but activity could be elicited by negative pipette pressure (suction). The conductance at zero membrane potential was 34 ± 2 pS ($n = 22$) for all combinations of 112 mM Na^+ or K^+ on the inside and outside. The channel does not conduct choline (selectivity ratio Na^+ :choline is greater than 5). When internal NaCl was replaced by sodium gluconate, the I-V curve was unchanged, excluding the possibility of a Cl^- channel. Similar conductances have been reported for other cation channels^{11,16-20}.

The relationship between channel activity and pipette pressure is shown in Fig. 2. Activation did not significantly depend on electrode solution. Channel activity was inherently unstable: the open probability changed spontaneously after being constant for minutes (Fig. 2). The pressure required to obtain an open probability of 0.1 was -1.75 ± 5 mm Hg (mean $\pm$ s.d., $n = 8$, range -7.0 to 7.4) for cell-attached patches and larger for inside-out patches (-5, -10 and -8.4 mmHg for three patches). There was no potential dependence of channel gating at negative potentials, but in some patches increased activity was observed at positive membrane potentials. Analysis of the kinetics of attached patches showed one open state with a dwell-time of 5.6 ± 0.9 ms (mean $\pm$ s.e.m., $n = 13$) and three closed states with dwell times of ~1 ms (1 kHz filtering), of ~10 ms and of ~20 to 100 ms. At 75 mM external Ca^{2+} , additional brief closures to the shut state were seen (Fig. 1). Stretch-activation produced increased open probability mainly by decreasing the long shut time. The spon-

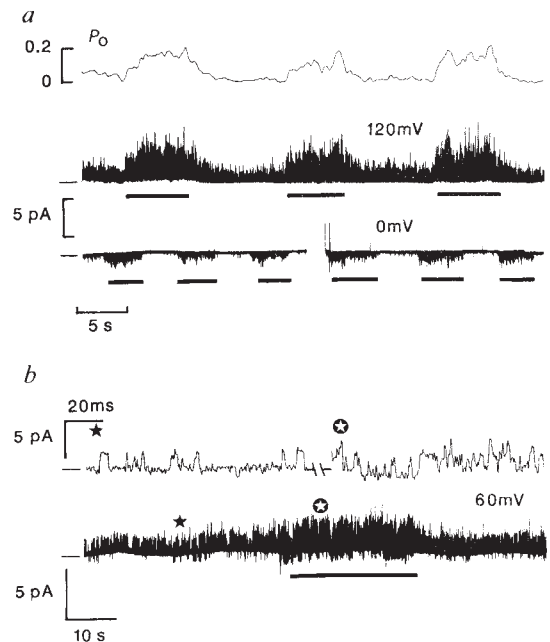


Fig. 3 Increase in the activity (=open probability) of the large (~200 pS) Ca^{2+} - and voltage-activated K^+ channel in two cell-attached patches caused by Ca^{2+} -injection through the stretch-activated channel in the same patch. *a*, With 75 mM CaCl_2 in the electrode. Top trace, open probability for the large K^+ channel at 120 mV BP (40 mV MP). Middle trace, actual single-channel K current from cell into pipette. Bar, pipette suction (~7 mmHg), which opens the SACs. At this high value of external Ca^{2+} , the outwards K^+ current was reduced from the normal value of ~8 pA to 3 pA (ref. 24). Lower traces, controls of the Ca^{2+} -current from electrode into cell through the SACs at 0 mV BP (-80 mV MP), obtained before (left) and after (right) the middle trace respectively. *b*, With 15 mM CaCl_2 and 90 mM NaCl in the electrode. At 60 mV BP (-40 mV MP) outward K^+ currents through the large K^+ channel (upward channel openings) are seen simultaneously with inward Ca^{2+} and Na^+ current through the SAC during stretch-activation (downwards channel openings), shown in the lower trace. Upper part, single-channel currents on an expanded time scale at indicated points (1 kHz filtering). The unidirectional Ca^{2+} current into the cell during channel opening is calculated from the constant field equation to be 0.06 pA in *a* and 0.17 pA in *b*.

taneous changes in open probability were also due to changes in the long shut time. Similar findings have been reported for other SACs^{10,11,21,22}. Internal Ca^{2+} was not necessary for channel activation. This was tested on seven inside-out patches. There is considerable evidence that SACs are attached to cytoskeletal strands^{10,22}. Deformation of the cytoskeleton may therefore result in stresses which open the attached SACs. The spontaneous activity of the SAC in cell-attached patches is probably an artefact due to deformation of the cell during seal formation. A small opening of the channels would also lead to an unacceptable influx of Ca^{2+} .

Recordings from two cell-attached patches, which contained the large (~200 pS) Ca^{2+} - and potential-activated K^+ channel¹³ together with several SACs, are shown in Fig. 3. The pipettes contained different concentrations of Ca^{2+} . When the SACs were activated by negative pipette pressure, the activity of the K^+ channels increased concomitantly with the activity of the SACs. In attached patches without Ca^{2+} in the pipette, the K^+ channel activity was independent of pipette pressure, both with and without SACs in the patch. K^+ -channel activity can be used to estimate internal Ca^{2+} concentration (ref. 13). It thus appears that external Ca^{2+} can enter the cell through the SACs and activate the neighbouring K^+ channels in this model system.

The K^+ -channel activity in a cell-attached patch was monitored as the external medium was shifted to lower tonicity

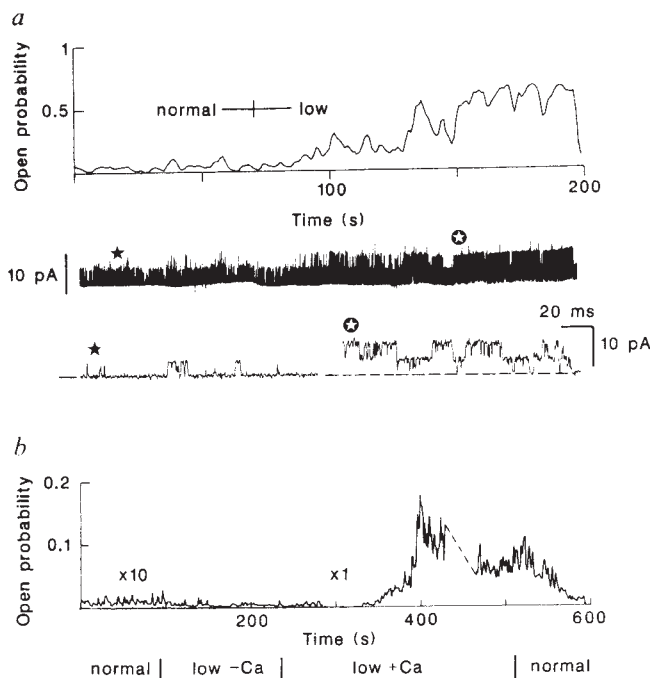


Fig. 4 The activity of the large K^+ channel monitored in cell-attached patches during change of external saline to low tonicity. *a*, With NaCl in the electrode, 100 mV BP (0 mV MP). Upper curve, open probability obtained from the patch current shown in the middle trace. Lower line, outwards single-channel currents at the times indicated on an expanded time-scale. Tonicity is shown as normal or low. *b*, With KCl in the electrode, 120–125 mV BP (30 mV MP). Tonicity shown thus; low-Ca, low tonicity without added Ca^{2+} but with 1 mM EGTA; low+Ca, low tonicity with 1 mM Ca^{2+} . Scale of recording shown as $\times 1$ or $\times 10$. To avoid changes in open probability due to changes of membrane potential at low tonicity (0 to 20 mV hyperpolarization), the bath potential was adjusted to obtain constant MP (dashed lines in *B*). Complete shift of external solution occurred within 45 s. Vertical bars, start of solution change.

Methods. Normal tonicity solution was usually NaCl saline and low-tonicity solution contained 56 mM NaCl, other ion concentrations being unchanged. In some experiments, the tissue was equilibrated first for several hours in NaCl saline to which was added 100 mM mannitol and the change to low tonicity was produced by removing mannitol. The two methods gave similar results²⁵.

(Fig. 4a). In six patches this caused increased activity of the large K^+ channel. When the medium is shifted to low tonicity without Ca^{2+} (Fig. 4b), the open probability remains unaltered until Ca^{2+} is added to the medium. It seems that external Ca^{2+} is necessary to produce increased K^+ -channel activity after hypotonic challenge. Thus either external Ca^{2+} enters the cell directly or is necessary for the release of Ca^{2+} from internal stores. Obviously Ca^{2+} might enter the cell directly through SACs which are activated by the initial cell swelling. The initial Ca^{2+} concentration was estimated from the open probability¹³ to be between 70 nM and 1.5 μ M. The latter high concentration could be caused by Ca^{2+} influx caused by membrane deformation during seal formation. In hypotonic medium, Ca^{2+} increased to 0.8–8.0 μ M. The membrane density of SACs was estimated to 0.2 μ m⁻² from the electrode resistance and the number of channels in the patch^{13,23}. Using a cell volume of 4,000 μ m³ and a surface of 1,000 μ m², the influx of external Ca^{2+} (1 mM) through 10% open SACs would cause internal Ca^{2+} to rise to 10 μ M in 100 s. This is of the same order of magnitude as the experimental results. The channel in the patch probably measures bulk Ca^{2+} . The Ca^{2+} level seen by the K^+ channels outside the patch could be higher, because they are close to the SACs. I suggest that the stretch-activated Ca^{2+}

channel is the transducing element for volume regulation through Ca^{2+} -activated K^+ channels.

This work was supported by a grant from the Danish Medical Research Council.

Received 10 September; accepted 18 September 1987.

1. Siebens, A. W. in *The Kidney: Physiology and Pathophysiology* (eds Seldin, D. W. & Giebisch, G.) 91–115 (Raven, New York, 1985).
2. Hoffman, E. K. *Fedn. Proc.* **44**, 2513–2519 (1985).
3. Kristensen, L. Ø. *Am. J. Physiol.* **251**, G575–G584 (1986).
4. Lauf, P. K. *Molec. Physiol.* **8**, 215–234 (1985).
5. Grinstein, S., Rothstein, A., Sarkadi, B. & Gelfand, E. W. *Am. J. Physiol.* **246**, C204–215 (1984).
6. Hoffmann, E. K., Simonsen, L. O. & Lambert, I. H. *J. Membrane Biol.* **78**, 211–222 (1984).
7. Larson, M. & Spring, K. R. *J. Membrane Biol.* **81**, 219–232 (1984).
8. Davis, C. W. & Finn, A. L. *Fedn. Proc.* **44**, 2520–2525 (1985).
9. Hoffmann, E. K., Lambert, I. H. & Simonsen, L. O. *J. Membrane Biol.* **91**, 227–244 (1986).
10. Guharay, F. & Sachs, F. *J. Physiol., Lond.* **352**, 685–701 (1984).
11. Cooper, K. E., Tang, J. M., Rae, J. L. & Eisenberg, R. S. *J. Membrane Biol.* **93**, 259–269 (1986).
12. Lansman, J. B., Hallam, T. J. & Rink, T. J. *Nature* **325**, 811–813 (1987).
13. Christensen, O. & Zeuthen, T. *Pflügers Arch. ges. Physiol.* **408**, 249–259 (1987).
14. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
15. Zeuthen, T., Christensen, O., Baerentsen, J. H. & laCour, M. *Pflügers Arch. ges. Physiol.* **408**, 260–266 (1987).
16. Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
17. Yellen, G. *Nature* **296**, 357–359 (1982).
18. Maruyama, Y. & Petersen, O. H. *Nature* **299**, 159–161 (1982).
19. Maruyama, Y. & Petersen, O. H. *Nature* **300**, 61–63 (1982).
20. Maruyama, Y. & Petersen, O. H. *J. Membrane Biol.* **81**, 83–87 (1984).
21. Guharay, F. & Sachs, F. *J. Physiol., Lond.* **363**, 119–134 (1985).
22. Sachs, F. *Fedn. Proc.* **46**, 12–16 (1987).
23. Sakmann, B. & Neher, E. in *Single-Channel Recording* (eds Sakmann, B. & Neher, E.) 37–51 (Plenum, New York, 1983).
24. Iwatzuki, N. & Petersen, O. H. *Biochim. biophys. Acta* **819**, 249–257 (1985).
25. Zeuthen, T. *J. Membrane Biol.* **66**, 109–121 (1982).

Retinal pigmented epithelial cells induced to transdifferentiate to neurons by laminin

T. A. Reh, T. Nagy & H. Gretton

Neuroscience Research Group and Department of Medical Physiology, University of Calgary, Calgary, Alberta, T2N 1N4 Canada

Although the regeneration of nervous tissue in the vertebrate is very limited, there are a few remarkable examples of this process¹. Understanding the factors that regulate CNS regeneration in those areas of the nervous system where it occurs, will doubtless provide generally applicable, essential information about the process. It has been known for some time that the amphibian retina regenerates following its destruction. Transplant studies^{2–4}, confirmed later by *in vitro* experiments⁵, have shown that one source of new neurons in regenerating retina is the retinal pigmented epithelium (RPE). RPE cells can transdifferentiate to either neurons or lens cells in culture^{5,6}, but little is known about the factors that regulate this process. A recent study *in vivo* of retinal regeneration provided evidence that the association of RPE cells with the retinal vascular membrane is an important step in transdifferentiation⁷. We report here that transdifferentiation *in vitro* is profoundly influenced by the substrate on which the cells are cultured; RPE cells plated on laminin-containing substrates frequently transdifferentiate into neurons. In addition, we have found a high concentration of laminin in the *Rana* retinal vascular membrane. Therefore, we propose that retinal regeneration is initiated by changes in the composition of the extracellular matrix that RPE cells contact early in the process.

Retinal pigmented epithelial cells were obtained by dissection and subsequent enzymatic dissociation of RPE from the posterior two-thirds of the eyes of *Rana catesbeiana* tadpoles. After one week of primary culture on a plastic substrate, the RPE cells were subcultured in multi-well plates on one of the following substrates: laminin; Englebreth-Holm-Swarm (EHS) sar-

Table 1 Growth of RPE cells on various substrates

Substrate	Neuronal clusters (% positive wells)	No. of γ -crystallin positive cells per well	^3H -thymidine incorporation d.p.m. per well
Plastic	10.5% (19)	146 $\pm$ 79 (6)	432.4 $\pm$ 224 (3)
Polylysine	5.2% (19)	—	326.5 $\pm$ 22.3 (3)
Collagen I	8% (27)	172 $\pm$ 80 (6)	501.1 $\pm$ 227 (3)
Fibronectin	0% (15)	—	413 $\pm$ 108.8 (3)
Laminin	86% (26)	1 (6)	268 $\pm$ 60.5 (6)
EHS membrane	94% (27)	7 $\pm$ 5 (6)	217.5 $\pm$ 79.9 (6)

Values are means of the number of measurements in parentheses, $\pm$ s.e.m. where appropriate. Neuronal clusters were identified as aggregates of small phase bright cells, bearing long processes (~ 10 times the cell diameter) that also bind with the neuron-specific 2D3 monoclonal antibody or tetanus toxin. Although neuronal clusters with neuritic processes were never observed on control cultures on plastic substrates, wells were considered positive if they contained any 2D3 staining cells. For ^3H -thymidine incorporation, cells were grown for 2 weeks *in vitro*, and $1 \mu\text{l ml}^{-1}$ isotope (20 Ci mmol^{-1}) added 48 h before trichloroacetic acid (TCA) precipitation. Cells were washed in normal frog Ringer's (three times), incubated for 10 min in $100 \mu\text{l}$ per well distilled water, and frozen and thawed to lyse the cells. Lysate ($50 \mu\text{l}$) was applied to Whatman No. 1 filter paper squares. The paper was washed three times in $200 \text{ ml } 10\% \text{ TCA } (4^\circ\text{C})$ then washed in ethanol and acetone. Scintillation fluid was added and the mixture was counted on an LKB β -counter.

coma basement membrane; collagen, type 1; fibronectin; polylysine; or plastic. The cells were examined daily for morphological changes, and stained with 2D3 or tetanus toxin at 3 to 12 weeks after subculture.

When subcultured on plastic, the RPE cells gradually became less pigmented; however, after six weeks *in vitro*, most cells still possessed some pigment, and frequently re-established the cuboidal array present in the primary culture. Many wells also developed structures similar to the "lentoid bodies" previously reported in cultures of newt and chick RPE (refs 6, 8 and 9), and characterized by clusters of elongated depigmented cells. The control cultures developed other morphologically distinct cell types, but no cells resembling neurons were observed. On two occasions (out of 19), we observed colonies of lightly pigmented cuboidal cells that stained with the 2D3 antibody in long-term (3-month) high-density control cultures, which may have represented early neuronal precursors like those found *in vivo* at peripheral retinal margin. No other evidence for a neuronal phenotype was observed. Although neuronal transdifferentiation has been reported in dissociated newt RPE cells cultured on plastic substrates, it developed in only a very few cases ($\sim 0.3\%$, ref. 5); as a result, it is not surprising that we failed to observe any neuronal-like, tetanus toxin/2D3 positive, cells in our control cultures.

When cultured on polylysine, collagen type I or fibronectin, the cells followed much the same course as cells plated on plastic. Occasionally, 2D3-positive cells were observed, but no other morphological features of neuronal cells were seen. But when the cells were plated on laminin, the cells underwent remarkable morphological changes. Within one to three days after subculture on laminin the RPE cells began streaming towards one another to aggregate into clusters. Initially, these aggregates took the form of networks (Fig. 1b) which frequently went on to form spherical clusters. At this time there was extensive depigmentation, and the cells formed compact spheres of small cells by 3–4 weeks. Soon after this, the cells extended fine processes resembling neurites, with growth cones at their tips (Fig. 1c). Indirect immunofluorescent staining of these cultures using either tetanus toxin or the 2D3 monoclonal antibody, showed that these clusters were positive for both markers (Fig. 1d), indicating that the RPE cells were now expressing a

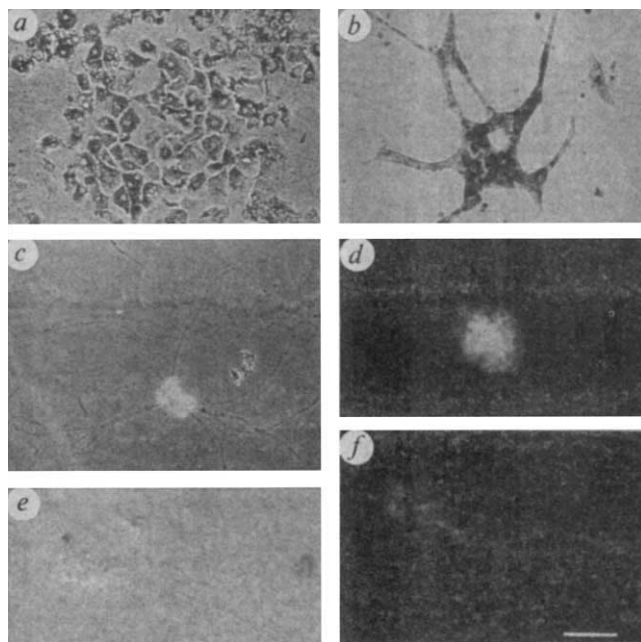


Fig. 1 Photomicrographs of the stages of transdifferentiation of *Rana catesbeiana* RPE cells when plated on laminin. *a*, Primary culture of RPE cells, just before they were subcultured. *b*, Network-like aggregate of RPE cells, two days after subculture on laminin ($100 \mu\text{g ml}^{-1}$). *c*, Neuronal cluster, six weeks after subculture on laminin ($100 \mu\text{g ml}^{-1}$). Note the fine, neuritic processes emanating from the cluster of small phase-bright cells. *d*, Similar neuronal cluster, now stained with the neuron-specific monoclonal antibody, 2D3. Scale bar, in *a*, *b* and *c* $100 \mu\text{m}$. *e*, Hoffman interference optics image and *f*, corresponding fluorescent micrograph, showing neurofilament staining of a neuronal cluster 8 weeks after subculture on EHS basement membrane.

Methods. For primary cultures, ten eyes were removed from midlarval stage *Rana catesbeiana* tadpoles, and the sclera, lens and retina dissected from the pigment in sterile Ringers' solution. Only the posterior two-thirds of the RPE was used; anterior RPE and iris were excluded. The RPE was treated with collagenase ($200 \text{ units ml}^{-1}$) for 15 min followed by 2.5 min in a 0.1% trypsin solution. The cells were then triturated, suspended in 10 ml of a 50% L-15 medium (with 10% serum and HEPES buffer added), plated on plastic tissue culture dishes (60 mm), and grown for three to five days. We ensured that no neurons, or germinal neuroepithelial cells were present in the primary cultures by examination by contrast phase microscopy or by indirect immunofluorescent staining with the neuron specific monoclonal antibody 2D3 (ref. 7). Counts of all cells in low-density cultures showed that 99.4% of the cells in the primary cultures were RPE cells; the remaining 0.6% of the cells were either choroid pigmented cells or fibroblasts; primary cultures that contained any 2D3 positive cells were not used in the experiments. For subculturing, the primary cultures were rinsed twice with EDTA Ringer's solution, and incubated in EDTA-Ringers with $100 \text{ units ml}^{-1}$ of collagenase (Sigma). Cells were then resuspended in media and plated on Falcon 3047 24-well plates, which had been previously coated with one of the following: $10\text{--}100 \mu\text{g ml}^{-1}$ laminin; $10\text{--}100 \mu\text{g ml}^{-1}$ fibronectin; $10\text{--}100 \mu\text{g ml}^{-1}$ polylysine; 0.1 mg ml^{-1} collagen. For immunohistochemistry, wells were incubated with tetanus toxin ($20 \mu\text{g ml}^{-1}$) followed by a rabbit anti-toxin serum (both gifts from Wellcome Research Labs) and Tetramethylrhodamine isothiocyanate (TRITC) conjugated goat anti-rabbit antiserum (Sigma). Alternatively, wells were incubated with the 2D3 monoclonal antibody (hybridoma culture supernatant or DEAE-cellulose purified from ascites) or a monoclonal antibody directed against the neurofilament subunit of relative molecular mass $200,000$ (M_r 200K) (Boehringer), and subsequently with a goat anti-mouse TRITC-conjugate. Wells were screened for neuronal cells both by phase microscopy and for 2D3 or tetanus toxin binding. The 2D3 antibody is a monoclonal, raised in our laboratory (by standard methods) to larval frog CNS membranes. Immunoblot analysis indicates that the antibody recognizes a glycoprotein of M_r $180\text{--}200\text{K}$, similar to N-CAM, that is specific to Ranids. The antigen is restricted to the germinal neuroepithelium and differentiated neural tissue at all embryonic and larval stages⁷. The neural specificity for frog NCAM has also been demonstrated in *Xenopus*^{26,27}. To detect the presence of RPE to lens transdifferentiation, anticycrystallin antiserum raised in rabbit against newt γ -crystallin (gift from D.S. McDevitt, University of Pennsylvania) was used to stain cultures, followed by a goat anti-rabbit FITC conjugate.

neuronal phenotype. Also, about one-third of the cells in these aggregates, as well as all of the neuritic processes, contained neurofilaments (Fig. 1e and f). The fact that only a fraction of the cells stain for neurofilament is interesting in light of reports that only retinal ganglion cells, and not other retinal cell types, normally possess neurofilaments *in vivo*¹⁰.

Another substrate that was particularly effective in promoting neuronal transdifferentiation was a commercial preparation of EHS basement membrane¹¹ (EHS-BM). Cells plated on this shared the same morphological changes that were observed for laminin substrates (Table 1): because EHS-BM contains considerable quantities of laminin¹¹, it is likely that the laminin was also the active component in this substrate.

From these results it seems that laminin-containing substrates induce dramatic morphological and biochemical changes in RPE cells *in vitro*, promoting their transdifferentiation to neurons. However, the laminin does not merely promote non-specific transdifferentiation of RPE cells, because transdifferentiation to lens cells seems to be suppressed on laminin or EHS-BM substrates: there were fewer lentoid bodies and staining of cultures with anti-crystallin antiserum showed a significant reduction in the number of γ -crystallin-containing cells per well (Table 1). Rather than having a direct effect, laminin may act by promoting proliferation of RPE cells. It has been proposed that RPE cells need to undergo several rounds of cell division before neural transdifferentiation *in vivo*^{5,12,13}. However, although ³H-thymidine incorporation and cell counts demonstrated cell proliferation in laminin- and EHS-basement-membrane-containing wells, there was also comparable cell proliferation in those wells with other substrates (Table 1). It is possible that we failed to identify neurons on non-laminin substrates, because they do not provide as favourable a surface for neurite outgrowth; however, polylysine and collagen support neurite outgrowth from tadpole retinal cells within one week after dissociation (T.R. and T.N., unpublished observations). Also the identification by tetanus toxin, 2D3 and neurofilament antibodies is not dependent on the presence of neurites (though the antibody staining is nearly always coincident with process-bearing aggregates of small phase-bright cells). Finally, it is also possible that the laminin exerts its effect by inducing tight clustering of the RPE cells. This is unlikely because clusters and aggregates frequently form on other substrates such as fibronectin and collagen, which do not promote transdifferentiation to a neuronal phenotype.

We have recently shown by an immunofluorescent analysis of regenerating *Rana* retina *in vivo*, that there is a high degree of association of migratory RPE cells with the retinal vascular membrane at the time the first new retinal neuroblasts appear⁷. Therefore, we proposed that contact with the retinal vascular membrane may be an important step in the process of retinal regeneration in amphibians. These results support and extend this hypothesis, by indicating that the relevant factor in the vascular membrane may be laminin. Laminin, a large glycoprotein present in basement membranes¹², influences many cellular processes including cell adhesion, growth, morphology and differentiation in a variety of epithelial cell types¹⁴. Analysis of the retinal vascular membrane using immunoblot and immunohistochemical analysis demonstrates the presence of several components normally found in basement membranes, including laminin, fibronectin, and heparan sulphate proteoglycan; however (Fig. 2), the relative amount of laminin in the vascular membrane greatly exceeds the amount found in the other tadpole basement membranes we examined, including Bruch's membrane, the basement membrane which RPE cells normally contact.

These results indicate that the process of retinal regeneration in amphibians is regulated, in part, by the extracellular matrix. Differences in the composition of Bruch's membrane and the vascular basement membrane are likely to be important in the control of this process; the small amounts of laminin in Bruch's

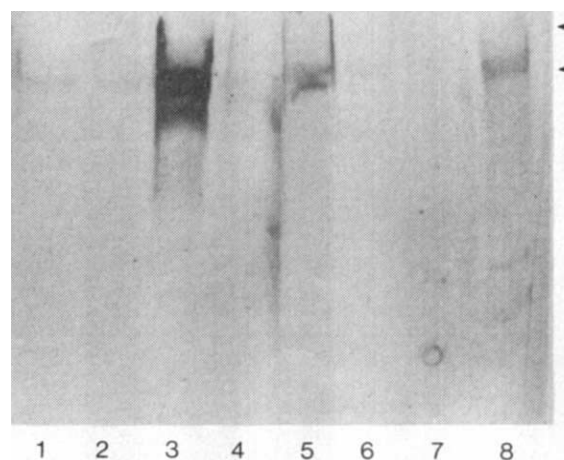


Fig. 2 Immunoblot of tadpole basement membranes with anti-laminin antisera. The blot shows the relative amounts of laminin immunoreactivity in sclera (lane 1), RPE (lanes 2 and 6), vascular membrane (lane 3), brain (lanes 4 and 7), lens (lane 5), and heart (lane 8). Arrowheads, positions of migration the A and B chains of laminin from mouse EHS-sarcoma; the A chain typically does not blot as well as the B chain. Similar blots with laminin antisera using two other extraction procedures gave comparable results; the vascular membrane contains far more laminin than the other basement membranes examined. Immunoblots run in parallel, but assayed instead for fibronectin, show a very different pattern of staining, with the highest levels found in Bruch's membrane and lens, and much less in the vascular membrane. Densitometry of the blots show a ratio of laminin/fibronectin of 1.8:1.0 for the vascular membrane and 0.3:1.0 for Bruch's membrane.

Methods. Basement membranes were dissected from *Rana* tadpoles, with the aid of a freeze-thaw regime and homogenized in a 0.5 M NaCl solution with phenyl-methylsulphonic acid (PMSF) and EDTA. The homogenate was centrifuged for 10 min at 10,000g and the pellet was resuspended and extracted by sonication for 12 h at 23 °C in 8 M urea, with 2% (v/v) mercaptoethanol, PMSF and 2% remaining tissue, and >80% of the laminin in the samples. The samples were then dialyzed for 24 h at 4 °C, protein concentration was assayed, (Biorad) and 10 μ g of each sample electrophoresed on a 5% SDS-polyacrylamide gel. Gels were then electroblotted, and incubated in anti-laminin (BRL) or anti-fibronectin (Calbiochem-Behring) antisera.

membrane are consistent with the maintenance of the RPE phenotype, whereas, when these cells come in contact with the higher laminin levels in the vascular membrane, the neuronal phenotype is favoured. Such basement membrane heterogeneity is involved in regulating the morphogenesis of several other developing tissues¹⁵⁻¹⁸, including mouse salivary gland branching and kidney nephron formation. Moreover, basement membranes obtained from various cell lines can profoundly influence the morphology and biochemistry, of several cell types *in vitro*, including kidney tubule epithelial cells¹⁹, corneal epithelial cells²⁰ and Sertoli cells²¹.

Our results also raise the possibility that extracellular matrix components, such as laminin, are involved in the inductive interactions that give rise to the retina and RPE during development. The results of experimental embryological manipulations indicate that the inductive signals are widespread in the embryo; contact with any mesenchymal cells is necessary for appropriate RPE development²², whereas contact with many other epithelial tissues, such as the ear vesicle, nasal placode or peritoneum can induce the developing RPE to become retina²³⁻²⁵. That changes in the composition of the extracellular matrix act as inductive signals in development is an emerging concept in the morphogenesis of many tissues in the embryo, and the results of the present study suggest this may be true for neuronal tissue as well.

We thank K. Radke for technical assistance and Drs D. Bowling and A. Bulloch for comments on the manuscript. T.A.R. is a Sloan Foundation Fellow and a Scholar of the Alberta Heritage Foundation for Medical Research. This research was also funded by the MRC of Canada.

Received 22 June; accepted 7 September 1987.

1. Jacobson, M. *Developmental Neurobiology* (Plenum, New York, 1978).
2. Stone, L. S. *Anat. Rec.* **106**, 89–109 (1950).
3. Stone, L. S. *J. exp. Zool.* **113**, 9–31 (1950).
4. Loposhev, G. V. & Sologub, A. A. *J. Embryol. exp. Morph.* **28**, 521–547 (1972).
5. Okada, T. S. *Curr. Topics dev. Biol.* **16**, 349–380 (1980).
6. Yamada, T. in *Cell Biology of the Eye* (ed. McDevitt, D. S.) 193–239 (Academic, New York, 1982).
7. Reh, T. A. & Nagy, T. *Dev. Biol.* **122**, 471–482 (1987).
8. Yasuda, K. *Dev. Biol.* **68**, 618–623 (1979).
9. Itoh, Y. & Eguchi, G. *Dev. Biol.* **115**, 353–362 (1986).
10. Cohen, J., Burne, J. F., Winter, J. & Bartlett, P. *Nature* **322**, 465–467.
11. Timpl, R. et al. *J. biol. Chem.* **254**, 9933–9945 (1979).
12. Eguchi, G., Abe, S. & Watanabe, K. *Proc. natn. Acad. Sci. U.S.A.* **71**, 5052–5054 (1974).
13. Eguchi, G. & Itoh, Y. *Pigment Epithelial Cell 1981: Phenotypic Expression in Pigment Cells* (ed. Seiji, M.) 271–278 (University of Tokyo Press, 1981).
14. Kleinman, H. K. et al. *J. Cell Biochem.* **27**, 317–325 (1985).
15. Ekblom, P., Lehtonen, E., Saxen, L. & Timpl, R. *J. Cell Biol.* **89**, 276–283 (1981).
16. Ekblom, P., Alitalo, K., Vaheri, A., Timpl, R. & Saxen, L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 485–489 (1980).
17. Ekblom, P. *J. Biol.* **91**, 1–10 (1983).
18. Banerjee, S. D., Cohn, R. H. & Bernfield, M. R. *J. Cell Biol.* **73**, 445–463 (1977).
19. Gospodarowicz, D., Lepine, J., Massoglia, S. & Wood, I. *J. Cell Biol.* **99**, 947–961 (1984).
20. Gospodarowicz, D. & Ill, C. R. *Expl Eye Res.* **31**, 181–199 (1980).
21. Hadley, M. A., Byers, S. W., Suarez-Quian, C. A., Kleinman, H. K. & Dym, M. *J. Cell Biol.* **101**, 1511–1522 (1985).
22. Holtfreter, J. *Arch. exp. Zellforsch.* **23**, 169–209 (1939).
23. Dragomirov, N. *Wilhelm Roux Arch. dev. Biol.* **129**, 522–560 (1933).
24. Dragomirov, N. *Wilhelm Roux Arch. dev. Biol.* **134**, 716–737 (1933).
25. Ikeda, Y. *Wilhelm Roux Arch. dev. Biol.* **136**, 676–680 (1937).
26. Jacobson, M. & Rutishauser, V. *Dev. Biol.* **116**, 524–531 (1986).
27. Kintner, C. R. & Melton, D. A. *Development* **99**, 311–326 (1987).

Cytotoxic T lymphocyte mediated lysis without release of serine esterase

Hanne L. Ostergaard*, Kevin P. Kane†, Matthew F. Mescher† & William R. Clark*‡

* Department of Biology and ‡ the Molecular Biology Institute, University of California Los Angeles, Los Angeles, California 90024, USA

† Division of Membrane Biology, Medical Biology Institute, La Jolla, California 92037, USA

Cloned cytotoxic T lymphocytes (CTL) can lyse target cells in a Ca^{2+} -dependent or Ca^{2+} -independent fashion, depending on the target cell used^{1,2}. We have used these Ca^{2+} -free assay conditions to determine whether there is serine esterase release in the absence of extracellular Ca^{2+} . We find that under conditions where there is significant target cell killing in the absence of Ca^{2+} , there is no detectable serine esterase release. To the extent that serine esterase and perforin-like molecules reside in the same granules, these results also suggest that target cell killing may occur in the absence of degranulation. Our results provide an example of target cell killing without serine esterase release, and the first indication that lysis and degranulation can be functionally dissociated.

The mechanism by which CTLs kill their appropriate target cells remains unclear. The prevailing view is that the CTLs release cytotoxic granules after a specific interaction with the target cell^{3,4}, probably through the T-cell receptor. These granules can be isolated from certain cloned cell lines and are able to lyse nucleated target cells^{5,6} in a manner analogous to complement^{7,8}. Cytotoxic granules contain, among other things, a pore forming protein called perforin, which can assemble in the membrane of the target cell, creating pores similar to those seen after complement-mediated lysis^{9–11}. Recently, a novel serine esterase has been identified in CTLs and other cytotoxic cells¹², which is released after specific target cell binding^{13,14}, and which has been localized to the cytotoxic granules^{15,16}.

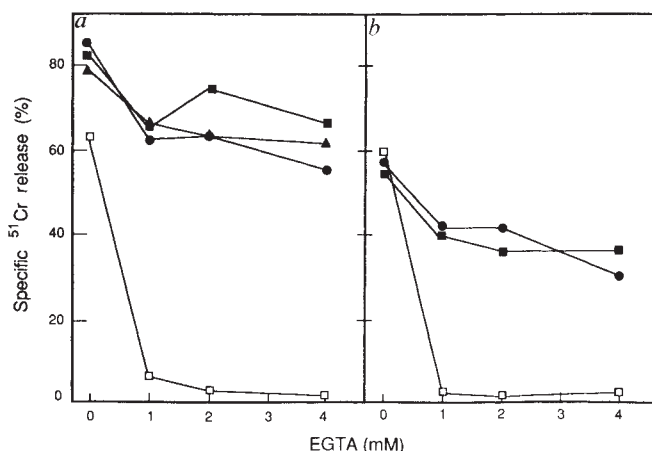


Fig. 1 Target cell lysis can proceed in the absence of Ca^{2+} . Ca^{2+} -dependent and -independent target cells were tested for their susceptibility to lysis by cloned CTL effector cells in the presence of varying concentrations of EGTA in *a*, direct cytotoxicity or *b*, lectin-dependent cytotoxicity assays. Symbols in *a*: $\blacktriangle$, AB.1 (H-2^d anti- H-2^b) vs EL-4 (H-2^b); $\bullet$, KB1.24 (H-2^k anti- H-2^b) vs EL-4; $\blacksquare$, c113 (H-2^k anti- H-2^b) vs EL-4; $\square$, c113 vs LB27.4 ($\text{H-2}^b \times \text{H-2}^d$); in *b*: $\circ$, 83.4 (H-2^b anti- H-2^d) vs EL-4; $\blacksquare$, KB1.24 vs R1.1 (H-2^k); $\square$, KB1.24 vs A20.2J (H-2^d).

Methods. The ^{51}Cr -labelled target cells 10^4 were mixed with 5×10^4 cloned CTL cells in a total volume of 200 μl in V-bottom 96-well plates. The assay medium contained 0.4 mM Ca^{2+} , and 4 mM Mg^{2+} , and 2% dialysed fetal calf serum (FCS). In *b*, the target cells were pulsed with 10 $\mu\text{g ml}^{-1}$ Con A for 30 min at 30 °C and washed twice before use. The assay plates were centrifuged at 300 g for 5 min before incubation for 3 h at 37 °C, after which the plates were again briefly centrifuged and 100 μl of supernatant collected and radioactivity counted. Data points are the average of triplicate samples. Spontaneous release was <15% of the total in all cases.

Exocytosis of the cytotoxic granules can thus be conveniently monitored by measurement of serine esterase activity released after specific target cell binding. In an examination of 15 CTL clones, with both allogeneic and hapten/self specificities, serine esterase release was observed in each clone on interaction with appropriate (but not inappropriate) target cells.

It has been accepted for a number of years that Ca^{2+} is required for cell-mediated lysis. Ca^{2+} is also required for assembly of perforin in the membrane of the target cell^{17,18}. Killing may occur in the absence of Ca^{2+} , however, and the requirement for Ca^{2+} is dictated by the particular target cell used¹. We have carried out a systematic analysis of the involvement of Ca^{2+} in CTL-mediated lysis, and have found, in agreement with ref. 1, that certain target cells can be killed in the complete absence of Ca^{2+} , even after exhaustive depletion of detectable stores of intracellular Ca^{2+} (ref. 2). The requirement for Ca^{2+} is clearly dictated by the target cell, because the same CTL clone can kill in either a Ca^{2+} -dependent or Ca^{2+} -independent fashion depending on the target cell used.

Figure 1*a* shows that CTL-mediated killing can occur in the presence of the Ca^{2+} -chelating agent EGTA, using the Ca^{2+} -independent target EL-4 (not all EL-4 sublines are Ca^{2+} -independent) and three different cloned CTL effector cell lines. At the concentrations of EGTA used, however, there is no killing of the Ca^{2+} -dependent target LB27.4, even though the same effector CTL clone (c113) was able to kill EL-4 target cells in the presence of EGTA. This Ca^{2+} -independence also extends to lectin-dependent cellular cytotoxicity (LDCC), in which specific target cell recognition is bypassed by use of an appropriate mitogenic lectin such as concanavalin A (Con A). There is significant lysis of the Ca^{2+} -independent targets EL-4 and R1.1 when they are pulsed with Con A, but not of the Ca^{2+} -dependent target A20.2J (Fig. 1*b*). Because our CTL clones remain 100% viable overnight in up to 4 mM EGTA (data not shown), we

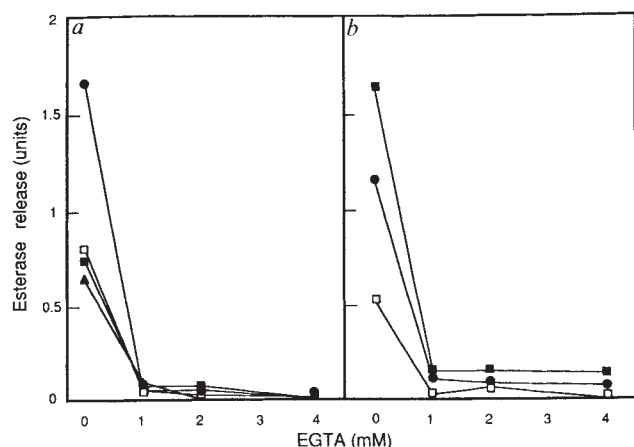


Fig. 2 Target cell lysis in the absence of Ca^{2+} does not involve secretion of BLT-serine esterase. Target cells were assayed for their ability to induce serine esterase release from CTL clones in the presence of different concentrations of EGTA in *a*, direct antigen-specific cytotoxicity or *b*, Con A-mediated cytotoxicity. Symbols in *a*: $\blacktriangle$, AB.1 vs EL-4; $\bullet$, KB1.24 vs EL-4; $\blacksquare$, cl13 vs EL-4; $\square$, cl13 vs LB27.4; in *b*: $\bullet$, 83.4 vs EL-4; $\blacksquare$, KB1.24 vs R1.1; $\square$, KB1.24 vs A20.2J.

Methods. For the assay, 10^5 CTL and 5×10^5 target cells were mixed in RPMI (0.4 mM Ca^{2+} , supplemented to a total of 4 mM Mg^{2+}) with 2% FCS in 100 μl total volume in flat-bottom 96-well plates. In *b*, the targets were first pulsed with 10 $\mu\text{l ml}^{-1}$ Con A for 30 min. The mixture was incubated for 3–4 h at 37°C after which 20 μl of supernatant was removed and assayed for serine esterase activity by addition to 180 μl of BLT substrate in the reaction buffer. The reaction mixture consisted of phosphate-buffered saline (PBS), pH 7.4 with 2×10^{-4} M *N*-benzyloxycarbonyl-L-lysine thiobenzyl ester and 1.1×10^{-4} M 5,5'-bithiobis (2-nitrobenzoic acid). The reaction was allowed to proceed for 20 min at 25°C, after which the optimal density was read at 412 nm wavelength in an enzyme-linked immunosolvent assay (ELISA) plate reader. Activity is expressed as optical density units per min per 10^6 CTL. EGTA did not inhibit serine esterase activity.

attribute the modest EGTA-dependent decrease in cytotoxicity of Ca^{2+} -independent target cells to the slight affinity of EGTA for Mg^{2+} . This would block cytotoxicity by preventing target cell binding.

We then tested whether serine esterase is released when there is killing in the absence of Ca^{2+} . When the antigen-specific target cell is used there is serine esterase release in the presence of Ca^{2+} (Fig. 2*a*). With even 1 mM EGTA, however, there is no detectable serine esterase release from any of the three CTL clones examined, even though there is significant killing of at least EL-4 cells when up to 4 mM EGTA is added to the assay (Fig. 1*a*). Antigen-specific target cells were pulsed with Con A and used to stimulate esterase release (Fig. 2*b*). When Ca^{2+} is present, there are high levels of esterase released, but there is no secretion when EGTA is added to the assay. Again it is clear that in the presence of EGTA, effector clones are able to kill Ca^{2+} -independent target cells (Fig. 1*b*), but this is not accompanied by release of detectable levels of serine esterase (Fig. 2*b*).

These data clearly show that CTL-mediated lysis can occur in the absence of detectable serine esterase release, and suggest that cytotoxicity, even by cloned interleukin-2 (IL-2)-driven CTL, may occur in the absence of effector cell degranulation. Evidence has been presented by others that serine esterase and perforin activity reside in the same cytoplasmic granule fraction^{15,16}. If this is true, our data also suggest that cloned CTL can kill target cells in the absence of degranulation. Other interpretations are possible. Although serine esterase and cytolytic molecules are found in the same granule fraction, they may not be found together in the same granule. A possible explanation of our data could then be selective degranulation, in which granules contain-

ing perforin are exocytosed, but those containing serine esterase are not. It is also possible that CTLs can kill by more than one mechanism. It has so far not been possible, for example, to correlate killing by primary CTLs with either cytotoxic granules or serine esterase activity^{19,20}, suggesting the existence of a killing pathway not associated with these functions. Although long-term-cloned, IL-2-driven CTLs may have gained a granule-associated killing mechanism, they may not necessarily have lost the non-granule, non-esterase mechanism of primary CTLs.

It is also possible that degranulation has nothing at all to do with cytotoxicity. In addition to target cell lysis, CTLs store and secrete a number of biologically important molecules such as serine esterases and γ -interferon. Release of these mediators on contact with a recognized target cell may be the primary function of degranulation. Target cell lysis by concentrated granule contents may not represent a normal physiological function. If, as has been claimed, exocytosed cytotoxicins and perforins are absolutely dependent on Ca^{2+} to lyse target cells^{5,17}, it would be difficult to show that such molecules are the lytic agents found in our experiments. The present results, with those of other authors, suggest that caution in interpreting the function of degranulation in IL-2-driven cell lines may be in order.

Received 16 June; accepted 1 September 1987.

1. Tirosh, R. & Berke, G. *Cell. Immun.* **95**, 113–123 (1985).
2. Ostergaard, H. L. & Clark, W. R. *J. Immun.* (in the press).
3. Young, J. D.-E. & Cohn, Z. A. *Cell* **46**, 641–642 (1986).
4. Henkart, P. A. *A. Rev. Immun.* **3**, 31–58 (1985).
5. Henkart, P. A., Millard, P. J., Reynolds, C. W. & Henkart, M. P. *J. exp. Med.* **160**, 75–93 (1984).
6. Podack, E. R. & Konigsberg, P. J. *J. exp. Med.* **160**, 696–710 (1984).
7. Podack, E. R. & Tschopp, J. *J. biol. Chem.* **257**, 15204–15212 (1982).
8. Podack, E. R. *J. cell. Biochem.* **30**, 133–138 (1986).
9. Dourmashkin, R. R., Deteix, P., Simone, C. B. & Henkart, P. A. *Clin. exp. Immun.* **42**, 554–560 (1980).
10. Dennert, G. & Podack, E. R. *J. exp. Med.* **157**, 1483–1495 (1983).
11. Podack, E. R., Young, J. D.-E. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8629–8632 (1985).
12. Pasternack, M. S. & Eisen, H. N. *Nature* **314**, 743–745 (1985).
13. Takayama, H. *et al. J. Immun.* **138**, 566–569 (1987).
14. Pasternack, M., Verret, C., Liu, M. & Eisen, H. *Nature* **322**, 740–742 (1986).
15. Masson, D., Nabholz, M., Estrade, C. & Tschopp, J. *EMBO J.* **5**, 1595–1601 (1986).
16. Young, J. D.-E. *et al. Cell* **47**, 183–194 (1986).
17. Young, J. D.-E., Leong, L. G., Liu, C.-C., Damiano, A. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5668–5672 (1986).
18. Young, J. D.-E., Podack, E. R. & Cohn, Z. A. *J. exp. Med.* **164**, 144–155 (1986).
19. Berke, G. & Rosen, D. *Transplant. Proc.* **19**, 412–416 (1987).
20. Anderson, C. G., Prochazka, G. & Dennert, G. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Exocytosis of cytolytic granules may not be required for target cell lysis by cytotoxic T-lymphocytes

Guido Trenn, Hajime Takayama & Michail V. Sitkovsky

Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Secretory processes have been implicated in the mechanism of target-cell lysis by cytotoxic T-lymphocytes (CTL) (refs 1, 2). Both CTL and cytotoxic large granular lymphocytes have cytolytic granules^{3,4}, containing the cytolytic molecules 'perforin'³ and 'cytotoxicin'⁴; perforin and cytotoxic lymphocytes can damage target cells through the membrane assembly of pores^{3–5}. The description of proteases which are cytotoxic cell-associated and granule-located has supported the granule exocytosis model of cytotoxicity mediated by cytotoxic lymphocytes, and has emphasized the similarities between cell-mediated and complement-mediated cytotoxicity^{6–14}. But recent experiments have¹⁵ challenged the importance of lytic granules and perforin in CTL activity^{15,16}. Here we report that CTL can be triggered to deliver a lethal hit

to target cells even when exocytosis of lytic granules has been abolished by removal of extracellular calcium. This dissociation of exocytosis of granules and delivery of the lethal hit suggests that cytolytic granules may not be involved in target-cell lysis by CTL.

Recent studies of granule exocytosis by CTL revealed that this process is regulated by lymphocyte antigen receptor (TcR) crosslinking, required extracellular calcium and involves protein kinase C activation¹⁷. TcR-triggered secretion of granule-associated enzymes β -glucuronidase¹⁷, serine esterase^{18,19} or a polypeptide^{17,18} labelled with [³H]diisopropylfluorophosphate was used to study exocytosis of cytolytic-containing granules. But situations have been reported^{17,20} in which significant target-cell lysis by CTL is accompanied by a very low or undetectable level of granule exocytosis. This suggests that although both exocytosis and 'lethal hit' delivery are triggered through T-cell receptor crosslinking there could be different molecular requirements for the execution of these two responses of CTL. Therefore, we designed experiments to test for a dissociation of exocytosis of lytic granules from target-cell cytotoxicity by taking advantage of the different Ca^{2+} requirements for exocytosis¹⁷ and target-cell lysis²¹.

Table 1 shows that granule exocytosis is triggered by crosslinking of TcR on the CTL surface with monoclonal antibody (mAb) raised against TcR, F23.1 mAb, which are chemically coupled to the proteins on the surface of the non-antigen-bearing target cells^{22,24}. The addition of di-magnesium EGTA (2mM) completely abolished exocytosis of granules, as shown by $0.3 \pm 0.7\%$ specific *N*- α -benzyloxycarbonyl-L-lysine thiobenzyl ester (BLT)-esterase release at 1:1 effector to target cell (E/T) ratio. Significant target-cell lysis ($30.6 \pm 2.1\%$) was still observed at this E/T ratio. Removal of extracellular Ca^{2+} also completely blocked TcR-triggered secretion of another CTL granule enzyme, β -glucuronidase (data not shown). These results confirm that extracellular Ca^{2+} is required for TcR-triggered exocytosis of CTL granules, consistent with our previous demonstration that extracellular Ca^{2+} is needed even for exocytosis triggered by the synergistic action of protein kinase C activators and Ca^{2+} ionophores¹⁷.

Binding of concanavalin A (Con A) to the surface glycoproteins of EL4 cells renders these cells susceptible to antigen-independent, CTL-induced lysis, perhaps from interaction of Con A with the T3-TcR complex on CTL (refs 25, 26) (Fig. 1). Significant target-cell lysis was observed even when EL4 cells pretreated with Con A (Con A-EL4) were mixed with CTL in the presence of di-magnesium EGTA. The effect of incubation with Con A-EL4 targets on granule exocytosis from CTL was monitored in parallel: Con A-EL4 cells were able to trigger granule enzyme secretion from CTL. This secretion was completely blocked in the presence of di-magnesium EGTA (Fig. 1). There was no exocytosis of granules observed in a specially prepared 'calcium-free' medium in the absence of EGTA although antigen bearing P815 target cells were still lysed by CTL (clone OE4), albeit less efficiently. We obtained similar results when EL4-Con A target cells were incubated with CTL OE4 in a 'calcium-free' medium (data not shown).

The same effects of EGTA on exocytosis and cytotoxicity were found with another CTL clone (130.17A), specific for H-2^d bearing target cells. In the presence of di-magnesium EGTA $15.1 \pm 1.1\%$ specific ⁵¹Cr-release could be observed from ⁵¹Cr-labelled P815 cells at a 1:1 E/T ratio. Parallel addition of di-magnesium EGTA inhibited granule exocytosis from $15.9 \pm 0.5\%$ to $1.38 \pm 2.5\%$ BLT-esterase secretion. Titration of the Ca^{2+} necessary for CTL-mediated target-cell lysis and for the target cell-induced, TcR-regulated exocytosis showed that at least $100 \mu\text{M}$ extracellular Ca^{2+} is required to reach or to achieve even a low level of exocytosis. Target-cell lysis, however, proceeds at extraordinarily low concentrations of Ca^{2+} , significant target cell death being observed even when 2 mM di-magnesium EGTA was added to the 'calcium-free' media.

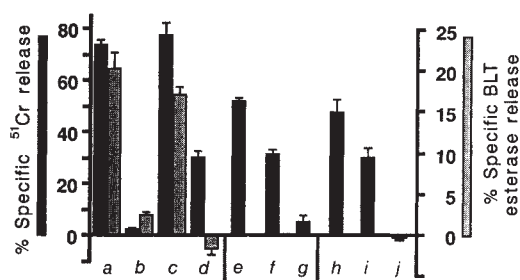


Fig. 1 Dissociation of exocytosis of granules and lethal hit triggering by CTL in a lectin-mediated, antigen-nonspecific cytotoxicity assay and in the antigen-specific cytotoxicity assay. Specific ⁵¹Cr release (dark shaded bars) and BLT-esterase secretion (light dotted bars) were measured after 6 h incubation. Bars a, b, c, d: Effect of extracellular Ca^{2+} removal on the target cell lysis and granule exocytosis by CTL. a, P815 were used as targets in the presence of Ca^{2+} ; b, EL4 were used as targets in the presence of Ca^{2+} ; c, EL4-Con A were used as targets in the presence of Ca^{2+} ; d, EL4-Con A were used as targets in the presence of 2 mM Mg_2EGTA . Bars e, f, g: Effect of anti-TcR mAb F23.1 on the lysis of EL4 in a 'calcium-free' medium. e, lysis of EL4-Con A in the presence of Ca^{2+} ; f, lysis of EL4-Con A in the 'calcium-free' medium; g, inhibition of EL4-Con A lysis in a 'calcium-free' medium by $5 \mu\text{g ml}^{-1}$ anti-TcR mAb F23.1 (ref. 23). Bars h, i, j: Effect of anti-LFA-1 mAb FD441.8 (ref. 29) on the lysis of antigen bearing P815 cells by CTL in the presence of 2 mM Mg_2EGTA . h, lysis of P815 in the presence of Ca^{2+} ; i, lysis of P815 in the presence of 2 mM Mg_2EGTA ; j, lysis of P815 in the presence of 2 mM Mg_2EGTA and $5 \mu\text{g ml}^{-1}$ anti-LFA-1 mAb.

Methods. CTL clone OE4 was incubated with either the antigen-bearing target cell P815 or with Con A pretreated non-antigen-bearing target cell EL-4 (EL4-Con A) in either RPMI-1640 medium with or without 2 mM Mg_2EGTA or in a calcium free DMEM medium (Biofluids) as indicated. Before incubation, cells were washed in calcium-free DMEM medium and resuspended in RPMI-1640 medium or in a calcium-free DMEM containing 5% FCS (dialysed against Dulbecco's magnesium- and calcium-free PBS) 2 mM MgCl_2 . After 6 h incubation at 37°C , serine esterase activity (BLT-esterase) in the supernatant were determined as a quantitative measure of the exocytosis of granules¹⁷. CTL clone OE4 was incubated with 5×10^4 P815 cells, EL4-Con A target cells per well. To prepare Con A-EL4 target cells EL-4 cells were labelled with ⁵¹Cr, washed and resuspended in $10 \mu\text{g ml}^{-1}$ concanavalin A-containing medium at 10^6 cells (refs 25, 26). EL4 cells with cell surface glycoproteins occupied by Con A (Ref. 25) were used as targets (EL4-Con A) in the presence or absence of Mg_2EGTA and/or different mAb in the incubation media. Control mAb 83-7-2 at $5 \mu\text{g ml}^{-1}$ did not affect either ⁵¹Cr release or exocytosis in TcR-triggered CTL. Spontaneous ⁵¹Cr release for EL4-Con A in the absence and presence of Mg_2EGTA after 6 h incubation was 13.1% and 14.9% respectively. Similar results were obtained at 2:1, 1:1 and 0.5:1 E/T ratios. Results at 1:1 E/T ratio are presented here.

Thus, by using 'calcium-free' medium or di-magnesium EGTA to decrease the Ca^{2+} concentration, we have found in three different assays of CTL-mediated cytotoxicity that it is possible to achieve target-cell lysis without detectable exocytosis of lytic granules from CTL. Target-cell lysis in the absence of Ca^{2+} requires occupation and crosslinking of the antigen-receptor, as shown in Table 1 and by the inhibition of cytotoxicity by soluble mAb to TcR on CTL (Fig. 1). We have also found that Ca^{2+} -independent cytotoxicity requires cell-to-cell contacts and that soluble mAb FD441.8 (ref. 29) to lymphocyte function-associated antigen (LFA-1) completely blocks CTL-induced target-cell lysis in the presence of di-magnesium EGTA (Fig. 1).

These data indicate that exocytosis of cytolytic granules from CTL cannot account for target cell lysis in a calcium-free medium, not only because no exocytic events occur in the absence of extracellular Ca^{2+} , but also because Ca^{2+} is required for pore formation by perforin^{3,28}. Our results do not completely

Table 1 Dissociation of exocytosis of granules and lethal hit triggering by CTL in a retargeting assay²²

Assay condition	% specific BLT-esterase release E/T ratio		% specific ⁵¹ Cr release E/T ratio	
	1/1	0.5/1	1/1	0.5/1
OE4 + P815	19.6 ± 0.7	33.2 ± 0.7	54 ± 3.1	32 ± 2.8
OE4 + EL4	0.7 ± 0.7	-0.03 ± 1.7	7.2 ± 0.3	3.0 ± 0.8
OE4 + EL4 = F23.1	15.2 ± 1.2	21.0 ± 1.1	68.8 ± 1.7	48.8 ± 2.4
OE4 + EL4 = F23.1 + EGTA	0.3 ± 0.7	2.1 ± 3.3	30.6 ± 2.1	12.3 ± 0.4

CTL clone OE4 (anti-H-2^d) (ref. 22) was incubated with the antigen-bearing target cell P815 (H-2^d) and non-antigen bearing target cell EL-4 (H-2^B). EL-4 was modified (EL4 = F23.1) by coupling it to the mAb F23.1 (ref. 23) which recognizes the TcR of OE4 CTL. The coupling procedure ('retargeting assay')^{22,24} uses the heterobifunctional cross-linking reagent *N*-succinimidyl-3-2-pyridyldithiol propionate (SPDP). SPDP in ethanol (50 µl, 20 mM) was added to 1 mg of affinity-purified F23.1 in 1 ml PBS (pH 7.2). After incubation for 30 min at room temperature, the modified mAb (SPDP-F23.1) was dialysed against PBS. To couple SPDP-F23.1 to the target cell, ⁵¹Cr-labelled EL-4 cells (10⁷) were washed twice with PBS, treated with 500 µM dithiothreitol for 30 min, washed with PBS and resuspended in 200 µl of SPDP-F23.1 at 0.5 mg ml⁻¹. After 30 min, the cells were washed, resuspended and adjusted to 10⁶ cells per ml. 5 × 10⁴ modified cells (EL4 = F23.1) were incubated with OE4 cells at an effector to target ratio of 1:1 and 0.5:1 either in medium (RPMI-1640, 5% fetal calf serum) or in medium with 2 mM ethyleneglycol bis (β-aminoethyl ether)-N,N,N',N'-tetra-acetic acid (EGTA) and 4 mM MgCl₂ (Mg₂EGTA). BLT was used for enzyme assay activity (BLT-esterase)^{18,19,30}. After 6 h incubation a specific release of ⁵¹Cr and secretion of BLT-esterase were estimated¹⁹. Both exocytosis and cytotoxicity measurements were performed on individual samples in a single experiment by incubating CTL clones with antigen bearing target cells or with irrelevant target cells which were chemically coupled to the anti-TcR mAb or pretreated with concanavalin A. Spontaneous ⁵¹Cr release (in the absence of CTL) for EL4 = F23.1 in the absence and presence of Mg₂ EGTA after 6 h incubation in this experiment was 16.2% and 25.3% respectively. Increased % exocytosis at lower E/T ratios both here and in the experiment described in Fig. 1 (in contrast to lower % ⁵¹Cr release) arise from calculation methods for BLT-esterase secretion^{17,19}. It is explained by the increased probability of CTL-target cell encounters and, subsequently, CTL activation (exocytosis triggering) when the number of targets exceeds the number of CTL.

exclude the involvement of exocytosis in 'lethal hit' delivery because vesicles containing cytolytic moieties other than perforin or markers as yet unknown, could be undetectable in the exocytosis assays used here. At least part of the cytolytic activity of these CTL could possibly be dependent on cytolytic granules as removal of extracellular Ca²⁺ does partly inhibit cytotoxicity (Table 1, Fig. 1). Two mechanisms for cytotoxicity could be used by CTL: one which is independent of extracellular Ca²⁺ and perforin, and probably used exclusively by peritoneal exudate CTL (ref. 15) and by *in vivo*-induced cytotoxic effector cells³¹ and another which is dependent on extracellular Ca²⁺ and perforin. The latter would probably operate only in those CTL clones expressing perforin and which are strongly inhibited by the removal of extracellular Ca²⁺. Future classifications of CTL clones could include these criteria. These data also indicate the existence of a response in cytotoxic T lymphocytes which is independent of extracellular Ca²⁺ and regulated by antigen-receptor binding. We are currently studying the TcR-triggered extracellular Ca²⁺-independent release of polypeptides as yet unidentified, into supernatants of metabolically labelled CTL.

The authors wish to thank Dr James Forman for providing CTL clone 130.17A, Dr William Clark for the EL4 cell line, Shirley Starnes for preparation of the manuscript and Drs Ronald Schwartz, Ethan Shevach, Ronald Germain and William Paul for discussion.

Received 6 July; accepted 17 September 1987.

- Bykovskaja, S. N., Rytenko, A. N., Rauschenbach, M. O. & Bykovsky, A. F. *Cell. Immun.* **40**, 175-185 (1978).
- Zagury, D. *Adv. exp. Med. Biol.* **14**, 149-160 (1982).
- Podack, E. R. *J. Cell. Biochem.* **30**, 133-170 (1985).
- Henkart, P. A. *Rev. Immun.* **3**, 31-58 (1985).
- Dourmashkin, R. R., Detreix, P., Simone, C. B. & Henkart, P. A. *Clin. exp. Immun.* **42**, 554-560 (1980).
- Reid, K. B. M. *Nature* **322**, 684-685 (1986).
- Pasternack, M. S. & Eisen, H. N. *Nature* **314**, 743-745 (1985).
- Kramer, M. D. *et al. J. Immun.* **136**, 4644-4651 (1986).
- Masson, D., Nabholz, M., Estrade, C. & Tschopp, J. *EMBO J.* **5**, 1595-1600 (1986).
- Masson, D. & Tschopp, J. *Cell* **49**, 679-685 (1987).
- Gershenfeld, H. K. & Weissman, I. L. *Science* **232**, 854-858 (1986).
- Brunet, J. F. *et al. Nature* **322**, 268-271 (1986).
- Lobe, C. G., Finlay, B. B., Paranchych, W., Paetkau, V. H. & Bleackley, R. C. *Science* **232**, 858-861 (1986).
- Brunet, J. F. *et al. J. Immun.* **138**, 4102-4105 (1987).
- Berke, G. & Rosen, D. *Trans. Proc.* **19**, 412-416 (1987).
- Golstein, P. *Nature* **327**, 12 (1987).
- Takayama, H. & Sitkovsky, M. *J. exp. Med.* **166**, 725-743 (1987).
- Pasternack, M., Verret, C. R., Liu, M. A. & Eisen, H. N. *Nature* **322**, 740-743 (1986).
- Takayama, H. *et al. J. Immun.* **138**, 566-569 (1987).
- Berrebí, G., Takayama, K. & Sitkovsky, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1364-1368 (1987).
- MacLennan, I. C. M., Gotch, F. M. & Golstein, P. *Immunology* **39**, 109-117 (1980).
- Staerz, U. D., Kanagawa, O. & Bevan, M. J. *Nature* **314**, 628-631 (1985).
- Staerz, U. D., Rammensee, H. G., Benedetto, J. D. & Bevan, M. J. *J. Immun.* **134**, 3994-4000 (1985).
- Krantz, O. M. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3375-3379 (1987).
- Sitkovsky, M. V., Pasternack, M. S., Lugo, J., Klein, J. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 573-577 (1984).
- Hubbard, S. C., Kranz, D. M., Longmore, G. D., Sitkovsky, M. V. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1852-1856 (1986).
- Tirosh, R. & Berke, G. *Cell. Immun.* **95**, 113-123 (1985).
- Ding-E Young, J., Damiano, A., DiNome, M. A., Leong, L. G. & Cohn, Z. A. *J. exp. Med.* **165**, 1371-1382 (1987).
- Dialynas, D., Loken, M., Sarmiento, M. & Fitch, F. W. *Adv. exp. Med. Biol.* **146**, 547 (1982).
- Takayama, H., Trenn, G. & Sitkovsky, M. V. *J. Immunol. Meth.* (in the press).
- Dennert, G., Anderson, C. G. & Prochazka, P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5004-5008 (1987).

Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase

Rob A. Gruters*, Jacques J. Neefjes†, Matthijs Tersmette*, Ruud E. Y. de Goede*, Abraham Tulp†, Han G. Huisman*, Frank Miedema*‡ & Hidde L. Ploegh†

* Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, Laboratory of Experimental and Clinical Immunology, University of Amsterdam, PO Box 9406, 1006 AK Amsterdam, The Netherlands

† The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis), Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands

Human immunodeficiency virus (HIV), the causative agent of AIDS, infects human lymphocytes and monocytes^{1,2}. An interaction between the viral envelope gp120 and CD4 protein is required to initiate an infectious cycle³⁻⁷. HIV infection *in vitro* induces syncytium formation by cell-to-cell fusion; this aspect of viral cytopathogenicity is even more dependent on gp120-CD4 interactions⁴. That gp120 is extremely heavily glycosylated (31-36 N-linked glycans per molecule), suggests involvement of N-linked glycans in the gp120-CD4 interaction⁸. We therefore investigated the effects of castanospermine, 1-deoxynojirimycin (dNM) and 1-deoxymannojirimycin (dMM), three trimming glucosidase inhibitors which perturb N-linked glycan structure⁹, on induction of the formation of syncytium between HIV-infected and CD4-expressing cells. The glucosidase inhibitors castanospermine and dNM, but not the mannosidase inhibitor dMM, inhibited syncytium formation and interfered with infectivity. The potential of glucosidase inhibitors as anti-HIV therapeutic agents deserves further investigation, especially because dNM and related compounds show little toxicity *in vitro* and *in vivo*⁹⁻¹¹.

‡ To whom correspondence should be addressed.

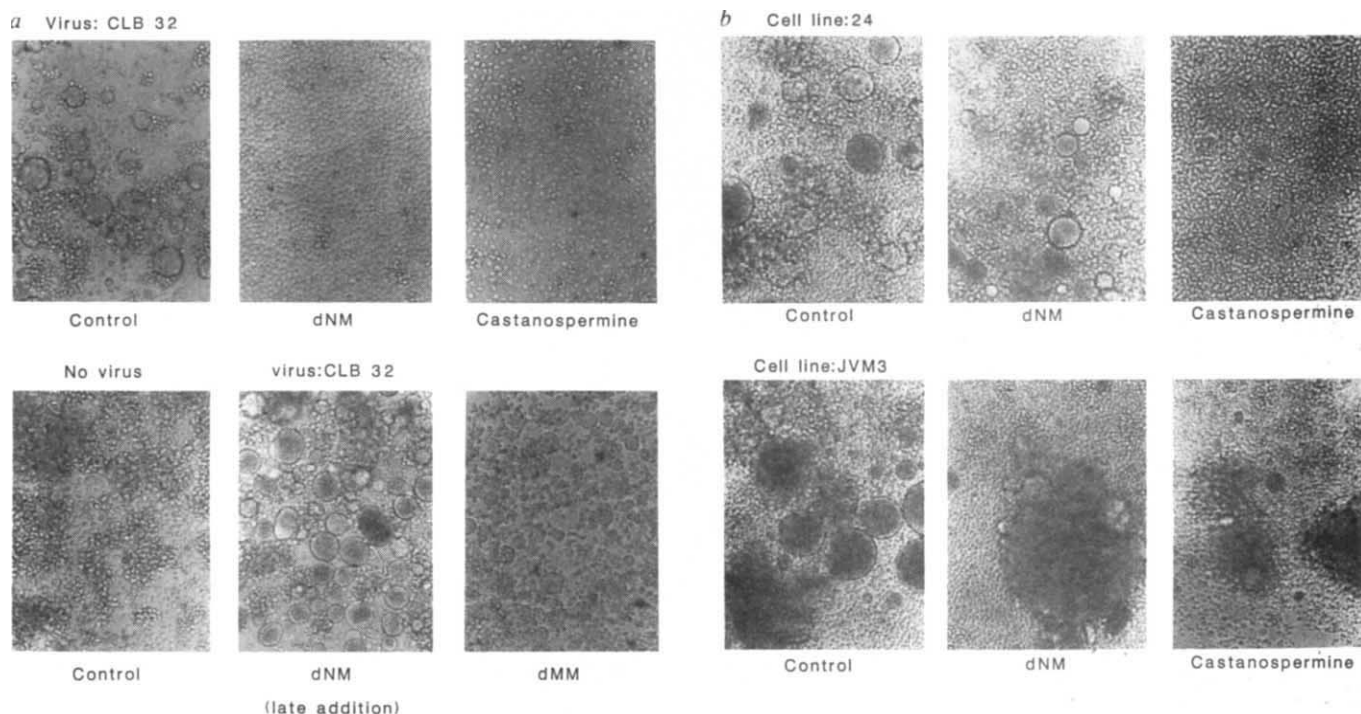


Fig. 1 *a*, Effect of glycosidase inhibitors on syncytium formation by HIV-infected promonocytic cells. The U937 cell line²⁰, uninfected (lower panel left) or infected with HIV isolate CLB-32 (upper panels and right-hand two lower panels), was grown in 1-ml cultures in the presence of 3 mM dNM, 1 mM castanospermine, 1 mM dMM or no inhibitor (as indicated). Cells were grown with inhibitor for at least three days in Iscove's modified Dulbecco's medium (IMDM) containing 10% fetal calf serum (FCS) in 1 ml cultures (Nunc, 24-well plates). Cells were collected and counted; 2×10^4 U937 cells were mixed with 1×10^5 non-infected H9 cells in wells of 96-well plates (Nunc). An experiment in which dNM was added at the time of mixing of U937 and H9 (dNM, late addition) shows that dNM itself does not perturb formation of syncytia. After 3, 6 and 24 h, cultures were scored for syncytia and photographed at the 3-h time point (magnification $\times 200$). Abbreviations: dNM, 1-deoxynojirimycin; dMM, 1-deoxymannojirimycin. *b*, Effect of glycosidase inhibitors on syncytium formation by HIV-infected B cells. The EBV-transformed B-cell lines 24 (ref. 18) (upper panels) and JVM3 (lower panels), infected with HTLV-IIIb prototype virus, were grown in the presence of 3 mM dNM, 1 mM castanospermine or no inhibitor (as indicated) in 1-ml cultures. Before counting, cell clusters were resuspended. Photographs shown were taken after 24 h. Note that after 24 h syncytia are visible in dNM-treated cultures, although to a lesser extent than in controls.

The human immunodeficiency virus (HIV) is the aetiological agent associated with AIDS (acquired immune deficiency syndrome), ARC (AIDS-related complex) and related diseases^{12,13}. HIV infects human cells that express the CD4 antigen, resulting in a marked depletion of CD4⁺ T cells in ARC and AIDS patients. Expression of the viral glycoprotein gp120 induces cell-to-cell fusion in cells that express large amounts of CD4. HIV-infected cells expressing small amounts of CD4 do not show cytopathic effects, but can still form syncytia with uninfected cells that express large amounts of CD4 (refs 1, 2 and 14), ultimately leading to the death of the latter cells.

The interaction between gp120 and CD4 is not understood. A non-glycosylated product of a recombinant gp120 gene or enzymatically deglycosylated gp120 were not able to inhibit HIV-induced cell fusion^{15,16}. The use of inhibitors of oligosaccharide processing⁹ such as the glucosidase I inhibitors 1-deoxynojirimycin (dNM) and castanospermine, and the mannosidase I inhibitor 1-deoxymannojirimycin (dMM) provides a more direct approach to studying the contribution of the sugar side chains to gp120-CD4 interactions. The predominantly high-mannose oligosaccharides carried by gp120 (ref. 8) could not be expected to change as a result of dMM treatment. However, glucosidase inhibitors should produce alterations in these high-mannose glycans, caused by retention of glucose residues.

As a simple model for studying syncytium formation, the CD4⁺ cell line H9 can be incubated with various HIV-infected cell lines expressing small amounts of CD4. This co-culture results in formation of giant multinucleated cells as soon as 3 h after mixing. This process can be blocked by antibodies against gp120 or OKT4A monoclonal antibodies^{1,17}. To study the effects of castanospermine, dNM and dMM on the process of syn-

cytium formation, the virus-producing cells were grown for at least 72 h in the presence of inhibitors. In such circumstances, all newly synthesized gp120 molecules are therefore aberrantly glycosylated, due to the presence of inhibitor. If the *N*-linked glycans were kept in the high mannose form by blocking mannosidase I with dMM, the extent of syncytium formation was similar to that seen in controls (Fig. 1). In contrast, syncytium formation by HIV-infected cells cultured with inhibitors of glucosidase I (dNM and castanospermine) was completely inhibited until 6 h after mixing, and almost completely blocked

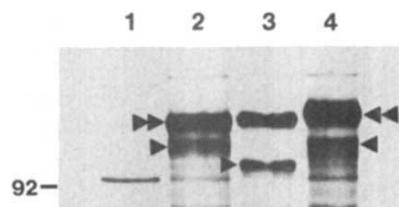


Fig. 2 Immunoprecipitation of lysates prepared from ³⁵S-methionine-labelled U937 cells, uninfected (lane 1) or infected with isolate CLB-32 in the presence of no inhibitor (lane 2), 1 mM dMM (lane 3) or 3 mM dNM (lane 4). Cells were labelled with ³⁵S-methionine for 6 h, and cell lysates were immunoprecipitated with the immunoglobulin G (IgG) fraction of a reference serum containing high-titre antibodies against gp120 as detected by immunoblot and immunoprecipitation on radio-iodinated virus preparations. The positions of gp160 (double arrow head), gp120 (single arrow head) and the marker for relative molecular mass *M_r* 92,000 (bar) are indicated. The incorporation of ³⁵S-methionine or ³H-thymidine was not notably changed in the presence of inhibitors.

Table 1 Effect of inhibitors of glycosylation on HIV infection

Target cells	Day after inoculation	Control		dNM		Castanospermine	
		RT assay (c.p.m. $\times 10^{-3}$)	Syncytia	RT assay (c.p.m. $\times 10^{-3}$)	Syncytia	RT assay (c.p.m. $\times 10^{-3}$)	Syncytia
H9 cells							
Experiment 1	14	12	>100*	<5†		0*	
Experiment 2	14	11	>100	<5	0		
PBL							
Donor							
1	15	120	238‡	120	35‡		
2	15	220	187	420	28		
3	10	120	193	310	54		
4	10	ND	198	110	123		
5	13	44	>200			200	46
6	13	14	60			51	30

RT, reverse transcriptase; ND, not determined. PBL of different donors were stimulated with PHA ($1 \mu\text{g ml}^{-1}$) and grown in IMDM, 10% FCS which contained 10 U ml^{-1} interleukin-2 (IL-2). After 2 days, 5×10^6 cells in 0.5 ml were incubated with or without inhibitor (3 mM dNM, 1 mM castanospermine) for 2 h, after which virus (100 TCID₅₀ units; for donors 5 and 6, 10 TCID₅₀ units) was added. After 1 h, incubation virus was washed away and cells were cultured in 5-ml cultures as indicated above. Donors 1 and 2 were supplemented with extra (5×10^6) phytohaemagglutinin (PHA)-stimulated, dNM-treated PBL on day 9. RT activity was determined as described¹⁹. The number of syncytia in PBL were counted in a random cross section of a 25-cm² flask (magnification $\times 100$). The experiment using H9 cells was similar, except that PHA stimulation was omitted and there was no IL-2 in the culture medium. H9 cells were grown in 1-ml cultures in a 24-well plate (Nunc) and the total number of syncytia was estimated.

* Total number in 1-ml culture estimated. † Not above background. ‡ Syncytia per cross section.

by castanospermine or markedly decreased by dNM for several days after mixing. The extent of inhibition of syncytium formation by dNM, and to a lesser extent that caused by castanospermine, was somewhat variable in these experiments and is difficult to quantitate. In Fig. 1 representative images of the effects of these inhibitors are presented.

The effects of dMM and dNM on the glycosylation of gp160, the precursor of gp120, was examined. Inhibition with dNM resulted in a precursor of slower mobility in SDS-PAGE (Fig. 2, lane 4), explained by the continued presence of glucose residues on the high-mannose oligosaccharides⁹. This inhibition of oligosaccharide processing did not seem to impair the cleavage of gp160 into gp41 and gp120, because roughly equal amounts of gp120 are observed for control and inhibitor-treated cultures (Fig. 2; compare lanes 2, 3 and 4). The glycans of gp160 are trimmed to the high-mannose form in both dMM-inhibited and control cells. After cleavage, the glycans in control cells are (partially) converted to the complex type; dMM blocks this process, resulting in a higher electrophoretic mobility for gp120 (Fig. 2, lane 3). As expected, uninfected cells do not produce gp160 or gp120 (Fig. 2, lane 1).

Figure 1a shows that addition of dNM (3 mM) or castanospermine (1 mM; not shown) at the time of mixing did not inhibit syncytium formation. When H9 cells were grown in the presence of the inhibitors and co-cultured with untreated, infected U937 cells, no inhibition of syncytium formation was observed, indicating that it is the alteration in glycan structure on gp120 which is responsible for the reduction in syncytium formation. A similar extent of blocking of syncytium formation by dNM and castanospermine was obtained with different Epstein Barr virus (EBV)-transformed B-cell lines including R, 24, 8-12 and JVM3, infected with HIV isolate HTLV-IIIb (ref. 18) (Fig. 1b). Radio-iodination of intact HIV-infected cells and subsequent analysis by SDS-polyacrylamide gel electrophoresis (PAGE), showed that gp120 was expressed on the cell surface if cells were grown in the presence of dNM (not shown). The glucosidase I inhibitors apparently had no effect on the amount of virus produced, based on assays for p24 and reverse transcriptase (RT) performed on sucrose-gradient-banded virus preparations (data not shown).

These inhibitors may have an even greater blocking potential under conditions which mimic more closely the situation *in vivo*. Such an experimental system would consist of exposure of freshly isolated lymphocytes to limited doses of virus, rather

than the use of permissive cell lines. Although virus production was increased compared to cultures of untreated peripheral blood lymphocytes (PBL), the numbers of syncytia were greatly reduced both in cultures of H9 cells and in five out of six cultures of normal lymphocytes that had been treated with dNM or castanospermine (Table 1). The increase in virus production in PBL grown in the presence of inhibitors might well be explained by the greater cell viability as a consequence of reduced syncytium formation. In H9 cell cultures the virus has clearly not established an infectious cycle in the presence of dNM.

Infectivity of HIV particles produced by H9 cells cultured in the continuous presence of dNM or castanospermine was reduced by a factor of ~ 100 (Table 2). This lesser inhibition of infectivity, as compared to the strong interference with syncytium formation, is compatible with the previous suggestion that of the two, cell fusion is dependent on stronger gp120-CD4 interaction^{4,15}.

Our results indicate that gp120 is synthesized in normal amounts and that virus production is not impaired in the presence of glucosidase I inhibitors. Viral infectivity and the formation of syncytia, on the other hand, are markedly reduced as a consequence of the continued presence of glucose residues on the N-linked glycans. The concentrations required to block syncytium formation and reduce infectivity are in the same range

Table 2 Infectious titres of HIV grown in the presence or absence of glucosidase inhibitors

Experiment	Control	dNM (3 mM)	Castanospermine (1 mM)
1	4.5	≤ 0.5	≤ 0.5
2	4.5	4.0	4.0
3	≥ 6.5	2.5	2.0
4	4.5	2.5	2.8
5	≥ 7.7	5.8	5.2

HIV-infected H9 cells were cultured for 4 days with or without inhibitor. Serial 10-fold dilutions of cell-free supernatant were added in quadruplicate to 5×10^5 H9 cells in a 24-well plate (experiments 1-3) or in octuplicate to 3×10^5 H9 cells in a 96-well microtitre plate (experiments 4 and 5). Infected cultures were scored after 14 days by p24 capture assay. Titres are expressed as the logarithm of TCID₅₀, the reciprocal of the dilution able to infect 50% of the cultures. Except for experiment 2, a reduction of at least 100-fold was observed with dNM or castanospermine treatment.

as those reported to inhibit glycan trimming⁹. The effects of castanospermine appear to be stronger, because even after more than 24 h syncytium formation is completely blocked. The reason for the difference between the inhibitory effects of dNM and castanospermine could be the faster metabolism of the former in our *in vitro* systems. It can not be excluded, however, that the dNM-imposed block is more leaky. Derivatives of dNM may well inhibit to a degree similar to castanospermine.

Thus, two structurally dissimilar glucosidase inhibitors, dNM and castanospermine, interfere with HIV infectivity and cell fusion, whereas the mannosidase I inhibitor dMM, the 2-epimer of dNM, does not. These findings argue in favour of our interpretation that the perturbed carbohydrate structure of gp120 or its precursor, imposed by blocking of the *N*-linked oligosaccharide trimming pathway, is responsible for the observed effects, rather than some non-specific activity in our inhibitor preparations or secondary effects of the inhibitors. These carbohydrate perturbations may result in a lower affinity of gp120 for the CD4 receptor.

Inhibitors of trimming glucosidase I may block or greatly decrease syncytium formation and viral infectivity. It has been observed that dNM and related compounds are hardly toxic in tissue culture⁹. In addition, derivatives of dNM have been used in humans in a clinical setting to control blood glucose levels, with no harmful side effects reported^{10,11}. It is not known whether castanospermine, dNM and related compounds offer a realistic prospect of a new therapy for those infected with HIV, but our data, obtained *in vitro*, warrant further exploration of this possibility.

We thank Drs H. Broxterman, G. van Marel and J. van Boom of the State University of Leiden, The Netherlands, for synthetic dNM, and Joline Claassen for technical assistance. This study was supported by the Netherlands Foundation for Preventive Medicine and the Netherlands Cancer Organization (Koningin Wilhelmina Fonds). F.M. is a senior fellow of The Royal Netherlands Academy of Arts and Sciences.

Received 5 June; accepted 22 September 1987.

1. Klatzmann, D. *et al.* *Nature* **312**, 767-768 (1984).
2. Nicholson, J. K. A., Cross, G. D., Callaway, C. S. & McDougal, J. S. *J. Immun.* **137**, 323-329 (1986).
3. Klatzmann, D. *et al.* *Science* **225**, 59-63 (1984).
4. Daigleish, A. G. *et al.* *Nature* **312**, 763-766 (1984).
5. Lifson, J. D., Rays, G. R., McGrath, M. S., Stein, B. A. & Engleman, E. G. *Science* **232**, 1123-1127 (1986).
6. Sodroski, J., Goh, W. C., Rosen, C. A., Campbell, K. & Haseltine, W. A. *Nature* **322**, 470-474 (1986).
7. McDougal, J. *et al.* *Science* **231**, 382-385 (1986).
8. Montagnier, L. *et al.* *Virology* **114**, 283-289 (1985).
9. Fuhrmann, U., Bause, E. & Ploegh, H. L. *Biochim. biophys. Acta* **825**, 95-110 (1985).
10. Joubert, P. H., Bam, W. J. & Manyane, M. *Eur. J. clin. Pharmac.* **30**, 253-255 (1986).
11. Shack, C., Röggl, G., Luger, A. & Scherthaner, G. *Eur. J. clin. Pharmac.* **30**, 417-419 (1986).
12. Barré-Sinoussi, F. *et al.* *Science* **220**, 868-871 (1983).
13. Popovic, M., Sarngadharan, M. C., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
14. Maddon, P. *et al.* *Cell* **47**, 333-348 (1986).
15. Matthews, T. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 5424-5428 (1987).
16. Putney, S. *et al.* *Science* **234**, 1392-1395 (1986).
17. Lifson, J. *et al.* *Nature* **323**, 725-728 (1986).
18. Tersmette, M., Miedema, F., Huisman, J. G., Goudsmit, J. & Melief, C. J. M. *Lancet* **i**, 815-816 (1985).
19. Tersmette, M. *et al.* *J. Immun.* (in the press).
20. Petit, A. J. C., Terpstra, F. G. & Miedema, F. *J. clin. Invest.* **79**, 1883-1889 (1987).

Phospho-dephospho-control by insulin is mimicked by a phospho-oligosaccharide in adipocytes

Susana Alemany*, Jose M. Mato* & Peter Strålfors†‡

* Departamento de Metabolismo, Nutrición y Hormonas, Fundación Jiménez Díaz, Avenida Reyes Católicos 2, 28040 Madrid, Spain

† Department of Physiological Chemistry, University of Lund, Box 94, S-22100 Lund, Sweden

The mechanism of insulin action is only partly understood^{1,2}. At one end of the signalling chain, the structure of the insulin receptor is known in detail³, and at the other end, insulin controls cellular metabolism by regulating the phosphorylation of serine and threonine residues in key target enzymes^{4,5}. The molecular events linking the occupied receptor to changes in target enzyme phosphorylation have remained obscure. Recently, insulin was shown to promote the hydrolysis of a phosphatidylinositol glycan with release of its polar head-group^{6,7}. The head group was reported to activate a high-affinity cyclic AMP-phosphodiesterase⁶ and pyruvate dehydrogenase⁸, to inhibit catecholamine-stimulated lipolysis⁹, and also to inhibit phospholipid methyltransferase¹⁰ and adenylate cyclase⁸. We report here that in intact adipocytes this head-group faithfully copies the insulin-directed effects on the phosphorylation and dephosphorylation of target proteins of the hormone.

The parent phospholipid is a 1-palmityl-2-palmitoyl phosphoglyceride with phosphodiester-linkage to a phospho-oligosaccharide head-group. The phospho-oligosaccharide (POS) consists of inositol monophosphate linked to non-acetylated glucosamine, and four residues of galactose with two additional phosphate groups that are not linked to the inositol moiety^{7,11,12}. The unusual *chiro* form of inositol is present¹², but its importance is questioned by work demonstrating the incorporation of radiolabelled *myo*-inositol into POS (ref. 7). This and reported differences in the fatty-acid composition^{7,11,13} could reflect struc-

tural differences in the glycopospholipid from different cells or tissues. A similar phosphatidylinositol glycan is involved in regulating the attachment of proteins to cell membranes¹⁴⁻¹⁶.

Insulin stimulates the phosphorylation of a protein of relative molecular mass (M_r) 116,000 (116K) in intact adipocytes (Fig. 1, lane *b*), identified as ATP citrate lyase¹⁷. POS mimicks this effect of insulin when added to intact cells (Fig. 1, lane *c*), an effect which is dose- (Fig. 2*a*) and time-dependent (Fig. 2*c*), being half-maximal at about 4 μ M and after 15 min. Parallel incubations showed that the half-maximal effect of insulin was at about 80 pM (Fig. 2*b*) and after 15 min (data not shown). Insulin also reduces the phosphorylation of an 84K phosphoprotein (Fig. 1), identified as hormone-sensitive lipase⁵. POS reproduces this dephosphorylation effect (Fig. 1) in a dose- (Fig. 2*a*) and time-dependent (Fig. 2*c*) process, which is half-maximal at about 4 μ M and after 5 min. Insulin shows a half-maximal effect at about 80 pM (Fig. 2*b*) and after 5 min (data not shown). At concentrations of insulin or POS giving maximal effects, their combined action does not further increase the response (data not shown). Treatment of POS with nitrous acid^{11,12} was found to abolish the insulin-mimicking effect (data not shown).

An increase in cAMP concentration causes the β -adrenergic agonist isoprenaline to increase the phosphorylation of a specific subset of phosphoproteins (Fig. 1, lane *d*). Insulin blocks this isoprenaline-induced increase in phosphorylation of the phosphoproteins with apparent M_r of 97K (glycogen phosphorylase¹⁸), 84K (hormone-sensitive lipase), 65K and 47K (Fig. 1, lane *e*). This effect of insulin is also reproduced by POS (Fig. 1, lane *f*), with the half-maximal effect being at about 3 μ M (Fig. 3), and that of insulin at about 50 pM. In Fig. 1, lanes *e* and *f*, both insulin and POS are seen to enhance the phosphorylation of a 40K protein, but this phosphorylation was not always obvious (compare with Fig. 1, lanes *b* and *c*).

In a cell-free system insulin has no effect on protein phosphorylation in the presence of [γ -³²P]ATP (Fig. 4, lane *b*), whereas POS markedly stimulates the phosphorylation of a 116K protein and of a 30K protein (Fig. 4, lane *c*), presumably ATP citrate lyase and the ribosomal protein S6 (ref. 19), respectively. POS causes a smaller increase or a reduction in the phosphorylation of other phosphoproteins (Fig. 4, lane *c*).

‡ To whom correspondence should be addressed.

Fig. 1 Effect of phospho-oligosaccharide on the state of phosphorylation of target proteins of insulin. Lane *a*, control with no additions; *b*, 1 nM insulin; *c*, 5 μ M phospho-oligosaccharide; *d*, 20 nM isoprenaline; *e*, 20 nM isoprenaline + 1 nM insulin; *f*, 20 nM isoprenaline + 6 μ M phospho-oligosaccharide; *g*, control with no additions. The positions of the following marker proteins are indicated; myosin (205 K), β -galactosidase (116 K), glycogen phosphorylase (97 K), bovine serum albumin (67 K), and ovalbumin (43 K). Also indicated are the following adipocyte phosphoproteins: 1, ATP citrate lyase; 2, glycogen phosphorylase; 3, hormone-sensitive lipase; 4, 65 K protein; 5, 47 K protein; 6, 40 K protein.

Methods. The phospho-oligosaccharide^{6,7,11,12} was purified from rat liver membranes^{9,11}; the procedure involved hydrolytic cleavage of a purified insulin-sensitive phosphoglyceride by the phosphatidylinositol-specific phospholipase C (ref. 23). The glycopospholipid isolated was free of contamination from all major phospholipids, including phosphoinositides^{11,12}. The presence of specific sugars was confirmed by gas chromatography and mass spectrometry¹². To quantify the phospho-oligosaccharide used, the amount of organic phosphate generated by phospholipase C treatment was determined²⁴. The concentration determination of the phospho-oligosaccharide assumed three phosphates per molecule¹². Adipocytes were isolated from the epididymal fat pads of 130–150-g Sprague-Dawley rats (Alab, Sweden) that were starved overnight and killed by cervical dislocation. Cells (final concentration 25 μ l packed cell-volume per ml) were incubated in Krebs-Ringer buffer (0.13 M NaCl, 5 mM KCl, 2.5 mM CaCl₂, 1.2 mM MgSO₄) containing 20 mM HEPES, pH 7.4, 3.5% (w/v) defatted bovine serum albumin, 100 nM phenylisopropyladenosine, 0.5 U ml⁻¹ adenosine deaminase, 5 mM glucose, 0.5 mM inorganic phosphate and 0.5 mCi ml⁻¹ of ³²P_i, at 37 °C for 2 h on a shaking waterbath²⁵. After 2 h incubation, all phosphoproteins had reached a steady state of ³²P incorporation. Cells were added to tubes containing the indicated final concentrations of additives and incubated for a further 20 min. Reactions were terminated by rapid flotation of the cells through a layer of silicone oil, thus separating the cells from the medium. Within 10 s of starting the centrifuge the cells were lysed in a solution containing SDS (40 mg ml⁻¹), ethanol, and inhibitors of proteinase (leupeptin, pepstatin, antipain at 20 μ g ml⁻¹), protein kinase and protein phosphatase (1.5 mM EDTA and 8 mM pyrophosphate), and frozen in an aluminium block at -20 °C (Strålfors and Honnör, manuscript in preparation). This procedure prevents the post-incubation modification of proteins. The total cell-protein was prepared for SDS-PAGE (D.S. and R. C. Honnör, manuscript in preparation) in 8% polyacrylamide gels²⁶. Gels were stained in 0.1% Kenacid blue (BDH), destained, dried between cellophane sheets and autoradiographed against Hyperfilm β -max (Amersham). Five different preparations of POS were used in these experiments which were repeated with similar results for different batches of POS and cells.

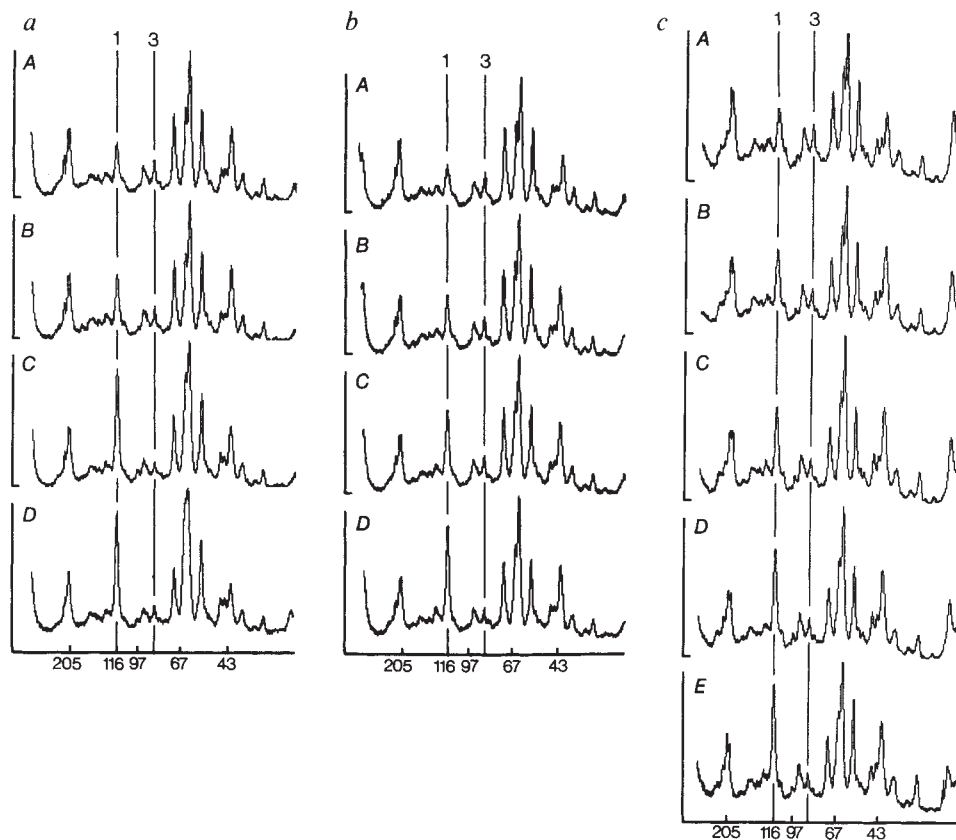
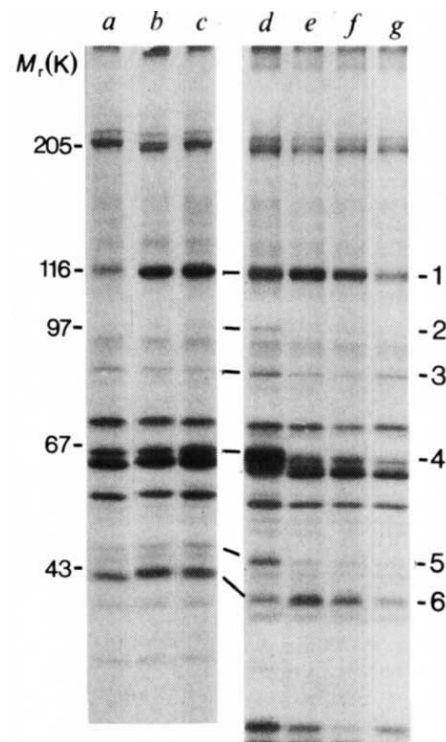


Fig. 2 Dose-response of phospho-oligosaccharide (*a*) and of insulin (*b*), and time-course of phospho-oligosaccharide (*c*) on adipocyte protein phosphorylation. *a*, Phospho-oligosaccharide at the following approximate concentrations: *A*, 1 μ M; *B*, 3 μ M; *C*, 5 μ M; *D*, 8 μ M. *b*, Control with no additions (*A*) and insulin at the following concentrations: *B*, 40 pM; *C*, 80 pM; *D*, 1,000 pM. *c*, Phospho-oligosaccharide at 10 μ M was incubated with the adipocytes for *A*, 0 min (no phospho-oligosaccharide, no change with time); *B*, 5 min; *C*, 10 min; *D*, 20 min; *E*, 40 min. Cells prepared and incubated with concentrations indicated of the phospho-oligosaccharide and insulin as in Fig. 1 legend, except that samples were withdrawn for SDS-PAGE after the incubation periods indicated. Gel autoradiographs of total cell-protein (legend to Fig. 1) were scanned with a Joyce-Loebl Chromoscan 3 densitometer, data digitized and corrected for protein-load in each lane of the gel: the protein load was determined by scanning densitometry and integration of the dried gel at 626 nm comparing with 205 K protein. The positions and M_r of the marker proteins described in the legend to Fig. 1 are indicated. Vertical lines: 1, ATP citrate lyase; 3, hormone-sensitive lipase.

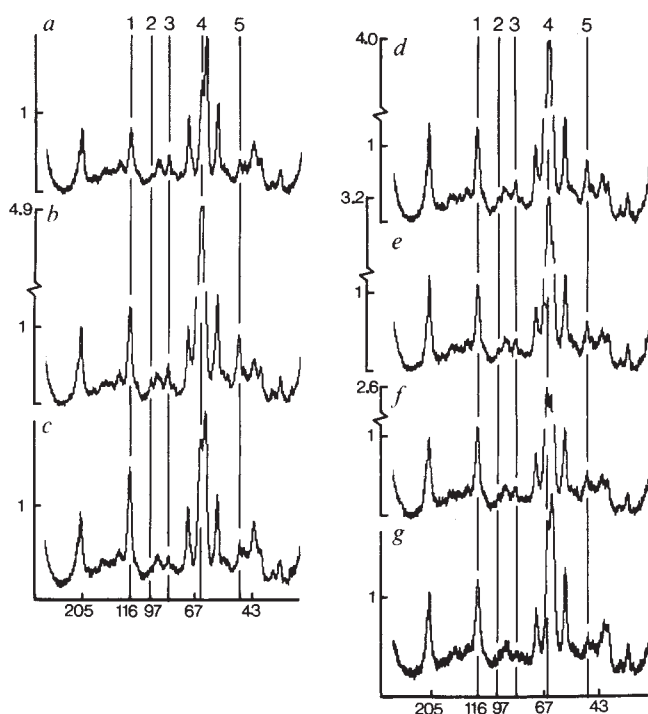


Fig. 3 Dose-response of phospho-oligosaccharide effect to block isoprenaline-stimulated adipocyte protein phosphorylation. Cells were prepared and incubated as described in legends Figs 1 and 2. *a*, Control with no additions; *b*, 20 nM isoprenaline; *c*, 20 nM isoprenaline plus 1 nM insulin; *d-g*, 20 nM isoprenaline plus the following concentrations of phospho-oligosaccharide: *d*, 2 μ M; *e*, 4 μ M; *f*, 7 μ M; *g*, 9 μ M. The broken scale on the ordinate is in arbitrary densitometric units. Vertical lines: 1, ATP citrate lyase; 2, glycogen phosphorylase; 3, hormone sensitive lipase; 4, 65 K protein; 5, 47 K protein.

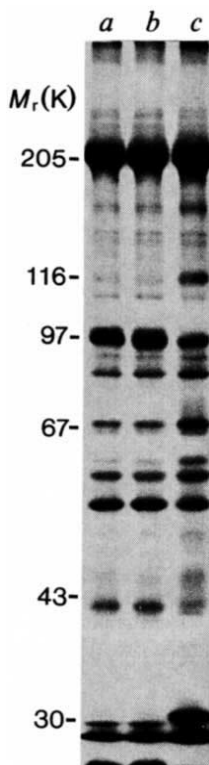
The phosphoprotein patterns obtained after SDS-polyacrylamide gel electrophoresis (PAGE) of whole adipocytes incubated with insulin effectively constitute a set of insulin-specific fingerprints which are copied in all detectable aspects by POS. Also, the relationship between the concentrations of POS and insulin needed for half-maximal effects was the same for the cAMP-dependent and cAMP-independent effects. The pronounced slow effect of insulin on the ATP citrate lyase phosphorylation was also reproduced by POS.

These results provide direct evidence that POS stimulates the same signalling pathways as insulin. The finding that both cAMP-dependent and cAMP-independent effects are copied by POS also indicates that hydrolytic production of POS must be an early event. The insulin-sensitive glucose transport in adipocytes is not affected by POS (ref. 9), indicating that POS is generated beyond a bifurcation in the pathways controlling glucose transport and intermediary metabolism. This agrees with previous reports that insulin controls the mobilization of the glucose transporters to the plasma membrane without affecting their phosphorylation²⁰. In these experiments, POS was added to intact cells, so it could have exerted its effects after being taken into the cells, or by interacting with the cell surface and then presumably with the insulin receptor. This last possibility is unlikely because although POS inhibited lipolysis, it did not affect glucose transport in fat cells⁹. The first possibility is favoured by our evidence that POS could stimulate phosphorylation of ATP citrate lyase and S6 in a cell-free extract of adipocytes.

The phospho-oligosaccharide differs from other proposed insulin-mediators in that it has a partially defined chemical structure, its generation is stimulated by insulin and it is active at physiologically relevant concentrations. The concentration required for half-maximal effects is comparable to that of the second messengers cAMP (ref. 21) and inositol-1,4,5-trisphosphate (ref. 22). It affects not only enzyme activity but also the phosphorylation status of insulin targets in the same way as insulin itself, as we have demonstrated both *in vivo* and in a cell-free system.

We thank Ulla Johannesson for technical assistance, Jose Puerta for phospho-oligosaccharide and Martin Low for inositol-specific phospholipase C. This work was supported by the A. Pahlsson Foundation, the M. Bergvall Foundation, the Crafoord Foundation, the Medical Faculty of the University of Lund, the Swedish Society of Medicine, and the Swedish Medical Research Council. S.A. held a FEBS Fellowship.

Fig. 4 Effects of phospho-oligosaccharide and insulin on protein phosphorylation in a cell-free adipocyte extract. Adipocytes were prepared as in Fig. 1 legend and preincubated for 90 min, washed in 50 mM HEPES, pH 7.3, 0.25 M sucrose, 100 nM phenylisopropyladenosine, 1 mM dithioerythritol, 5 mM benzamide, 10 μ g ml⁻¹ of pepstatin, antipain and leupeptin. The buffer was aspirated, and the cells broken by vortexing with 1 mm glass beads to give concentrated (20% v/v) cytosolic extract after 2 min centrifugation at 10,000 $\times$ g. This extract was incubated for 5 min at 30 °C with 0.7 mM [γ -³²P]ATP (3 mCi μ mol⁻¹), 2.5 mM MgCl₂ and the following additions: *a*, control with no addition; *b*, 100 pM insulin; *c*, 7 μ M phospho-oligosaccharide. Incubations were terminated by mixing with SDS-PAGE sample solution and boiling. Autoradiography as in Fig. 1.



Received 19 June; accepted 22 September 1987.

1. Czech, M. P. *Molecular Basis of Insulin Action* (Plenum, New York, 1985).
2. Belfrage, P., Donné, J. & Strålfors, P. *Mechanisms of Insulin Action* (Elsevier, Amsterdam, 1986).
3. Ullrich, A. et al. in *Mechanisms of Insulin Action* (eds Belfrage, P., Donné, J. & Strålfors, P.) 1-12 (Elsevier, Amsterdam, 1986).
4. Avruch, J., Nemenoff, R. A., Pierce, M., Kwok, Y. C. & Blackshear, P. J. in *Molecular Basis of Insulin Action* (ed. Czech, M.) 263-298 (Plenum, New York, 1985).
5. Strålfors, P., Björnell, P. & Belfrage, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3317-3321 (1984).
6. Saltiel, A. R. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5793-5797 (1986).
7. Saltiel, A. R., Fox, J. A., Sherline, P. & Cuatrecasas, P. *Science* **233**, 967-972 (1986).
8. Saltiel, A. R. *Endocrinology* **120**, 967-972 (1987).
9. Kelly, K. L., Mato, J. M., Merida I. & Jarett, L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6404-6407 (1987).
10. Kelly, K. L., Mato, J. M. & Jarett, L. *FEBS Lett.* **209**, 238-242 (1986).
11. Mato, J. M., Kelly, K. L., Abler, A. & Jarett, L. *J. biol. Chem.* **262**, 2131-2137 (1987).
12. Mato, J. M. et al. *Biochem. biophys. Res. Commun.* **146**, 764-770 (1987).
13. Saltiel, A. R., Sherline, P. & Fox, J. A. *J. biol. Chem.* **262**, 1116-1121 (1987).
14. Cross, G. A. M. *Cell* **48**, 179-181 (1987).
15. Ishihara, M., Fedarko, N. S. & Conrad, H. E. *J. biol. Chem.* **262**, 4708-4716 (1987).
16. Low, M. G. *Biochem. J.* **244**, 1-13 (1987).
17. Ramakrishna, S. & Benjamin, W. B. *J. biol. Chem.* **254**, 9232-9236 (1979).
18. Belfrage, P., Fredrikson, G., Nilsson, N. Ö. & Strålfors, P. *FEBS Lett.* **111**, 120-124 (1980).
19. Rosen, O. M., Rubin, C. S., Cobb, M. H. & Smith, C. J. in *Protein Phosphorylation* (eds Rosen, O. M. & Krebs, E. G.) 771-781 (Cold Spring Harbor, 1981).
20. Gibbs, E. M., Allard, W. J. & Lienhard, G. E. *J. biol. Chem.* **261**, 16597-16603 (1986).
21. Beavo, J. A., Bechtel, P. J. & Krebs, E. G. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3580-3583 (1974).
22. Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67-69 (1983).
23. Low, M. G. & Finean, J. B. *Biochem. J.* **167**, 281-284 (1977).
24. Leloir, L. F. & Cardini, C. E. *Meth. Enzym.* **3**, 843-844 (1957).
25. Honnor, R. C., Dhillon, G. S. & Londos, C. *J. biol. Chem.* **260**, 15122-15129 (1985).
26. Laemmli, U. K. *Nature* **277**, 680-685 (1970).

Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis

Karel Odink*, Nico Cerletti*, Josef Brügggen*, Roger G. Clerc*, Lajos Tarcsay*, Gabriele Zwadlot†, Gerold Gerhards†, Renate Schlegel† & Clemens Sorg†

* Department of Biotechnology, Ciba-Geigy, K681-442, 4002 Basel, Switzerland

† Institute of Experimental Dermatology, University of Münster, 4400 Münster, FRG

The aetiology and cellular mechanism of chronic inflammatory processes are poorly understood. Macrophages act prominently in the inflammatory response and we report here that they express two calcium-binding proteins. The expression of these proteins, referred to as MRP-8 and MRP-14, is specific for cells of myeloid origin, namely granulocytes, monocytes and macrophages, and is observed in blood granulocytes and monocytes but not in normal tissue macrophages. In acutely inflamed tissues, macrophages can express MRP-14 but not MRP-8, and in chronic inflammations, such as primary chronic polyarthritis, infiltrate macrophages express both MRP-8 and MRP-14. Characterization of MRP-8 and MRP-14 could therefore be useful to the understanding of cellular processes induced in chronic inflammation.

The proteins MRP-8 and MRP-14 were isolated from human peripheral blood mononuclear cell cultures as part of a complex, using a monoclonal antibody directed against human macrophage migration inhibitory factor (MIF). The monoclonal

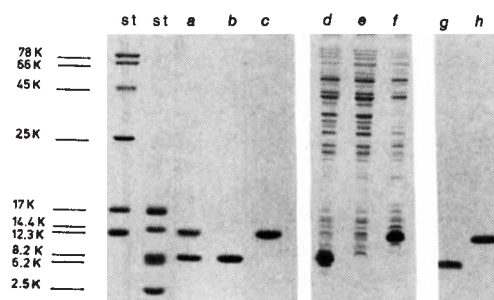


Fig. 1 SDS polyacrylamide gel electrophoresis (PAGE) analysis of natural and recombinant MRP-8 and MRP-14 under reducing conditions. Track *a*, eluate of 1C5-antibody column; *b*, HPLC-purified natural MRP-8; *c*, HPLC-purified natural MRP-14; *d*, *E. coli* extract containing recombinant MRP-8; *e*, control extract; *f*, *E. coli* extract containing recombinant MRP-14; *g*, purified recombinant MRP-8; *h*, purified recombinant MRP-14; *st*, protein standards, M_r s shown on left.

Methods. Recombinant MRP-8 and MRP-14 were expressed in *E. coli* from a *trp*-promoter as described⁹. For MRP-8 a restriction site at position 70, for MRP-14 at position 67, were used for the isolation of the major part of the coding sequence. With synthetic linkers the first amino acids were restored and the AUG was placed 12 nucleotides from the ribosome binding site. On tryptophan starvation *E. coli* cells containing the recombinant plasmids, pMRP-8-*trp* and pMRP-14-*trp* respectively, produced recombinant protein to between 15–30% of total protein. Purification of natural and recombinant MRP-8 and MRP-14 (N.C. *et al.*, in preparation) was as follows: culture supernatants of Con A-stimulated peripheral blood mononuclear cells were chromatographed over an 1C5-antibody affinity column^{1,8} and natural MRP-8 and MRP-14 were eluted at pH 2.5 and separated by reverse-phase HPLC (Vydac C4 column). The recombinant proteins were purified from crude *E. coli* lysates containing 8M urea, by DEAE-Trisacryl M and SP-Trisacryl M ion-exchange chromatography, followed by size exclusion HPLC (Shimadzu Shimpack Diol 150) or reverse phase HPLC (Nucleosil C18). Partial amino-acid-sequence analysis and mass spectrometry of the recombinant proteins confirmed the sequence in Fig. 2, including the N-terminal methionine.

antibody (1C5) used for the immuno-affinity purification binds MIF without neutralizing MIF activity¹. The relative molecular masses (M_r) of the proteins on SDS-PAGE are 8,000 and 14,000 (8 K and 14 K), respectively (Fig. 1, track *a*). Therefore we refer to these proteins as MIF related proteins, MRP-8 and MRP-14. Using oligonucleotide probes based on the partial amino-acid sequence we isolated complementary DNA clones for MRP-8 and MRP-14 (Fig. 2). The sequence of MRP-8 cDNA has an open reading frame (ORF) of 279 nucleotides, predicting a protein of 93 amino acids and M_r 10,835. The sequence of MRP-8 is identical to the sequence for the cystic fibrosis (CF) antigen², with one exception. We find an additional G residue in position 292, which shifts the reading frame for the last 15 amino acids. This difference could reflect a natural variation in alleles or could arise from a sequencing error. The sequence in Fig. 2 was confirmed for three independent cDNAs and for a chromosomal MRP-8 gene (E. Lagasse and R.G.C. *Molec. cell. Biol.*, submitted). In addition, the predicted mass of MRP-8 was confirmed by mass spectrometry of recombinant MRP-8. MRP-14 cDNA (Fig. 2*b*) contains an ORF of 352 nucleotides predicting a protein of 114 amino acids and M_r 13,242. MRP-8 and

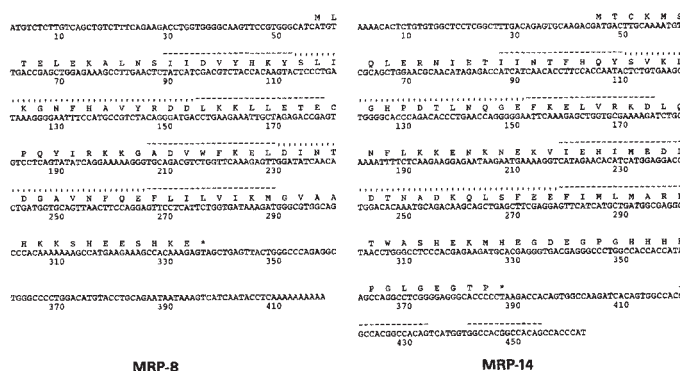


Fig. 2 Nucleotide and amino-acid sequence of MRP-8 and MRP-14. Stop codons are indicated by an *. Stretches of amino acids typical for Ca^{2+} -binding domains³ are indicated with ~~~ for the α -helices and ,,, for turns. Sequences overlined with ---- in MRP-14 indicate two major repeats containing twice the palindromic sequence GGCC. More of such repeating sequences were found in the 3'-non-translated region of the mRNA, as predicted from the chromosomal gene (E. Lagasse & R.G.C. *Molec. cell. Biol.*, submitted). The presence of these sequences could explain why the 3' end of the mRNA is not represented in any of the 5 clones analysed. Primer extension experiments have shown that the 5' end of the mRNAs map at position 1 in the figure.

Methods. Natural MRP-8 was purified from a 1C5 affinity column eluate (see Fig. 1) by SDS-PAGE and its N-terminal sequence determined up to residue 61 on a gas-phase sequencer. MRP-14 was purified as in Fig. 1 (track *C*) and appeared to be N-terminally blocked. After partial *S. aureus* V8 digestion three oligopeptides were purified by reverse phase HPLC and sequenced (residues 14–27, 40–59 and 7–94, N.C. *et al.*, in preparation). RNA was prepared from human blood leukocytes¹ as described¹⁰ with the following modifications: after lysis in guanidium thiocyanate most of the DNA was removed by three cycles of phenol extraction and phase separation with chloroform before the RNA could be precipitated through a CsCl-cushion. cDNA was prepared¹¹ using mRNA from a sucrose gradient fraction sedimenting at 9S as template, and cloned into a derivative of pUC9, from which the *lacZ*-promoter bearing *Hae*II-*Hind*III fragment was deleted. The use of this vector was necessary to maintain stability of MRP-14 cDNA inserts. mRNA and cDNA clones were detected by hybridization using mixed oligonucleotide probes which were deduced from amino-acid sequences 14–19 for MRP-8 and 14–22 for MRP-14. The MRP-8 probe was 5'-TAYTTRTGRTANACRTC-3', and MRP-14 probe 5'-TAYTGRTGAAIGTRTTIATINGT-3', where Y, R, I and N stand for pyrimidines, purines, inosine and one of the 4 common nucleotides. DNA sequences were determined¹² for at least two independent clones and in both strands.

Fig. 3 Western blot analysis of an SDS-polyacrylamide gel containing natural and recombinant MRP-8 and MRP-14. Tracks *a* and *d*, crude *E. coli* extract containing recombinant MRP-8 and MRP-14; tracks *b* and *e*, natural MRP-8 and MRP-14 (1C5 antibody eluate); tracks *c* and *f*, lysate of 5×10^5 adherent macrophages. Tracks *a-c* were developed with anti-MRP-8 serum, tracks *d-f* with anti-MRP-14 serum.

Methods. Antisera were raised in rabbits using recombinant MRP-8 and MRP-14. Anti-*E. coli* activities were removed by absorption. Immunocomplexed proteins were visualized using goat anti rabbit IgG immunoperoxidase. Macrophages were prepared from human peripheral blood by percoll density fractionation and adherence¹³.

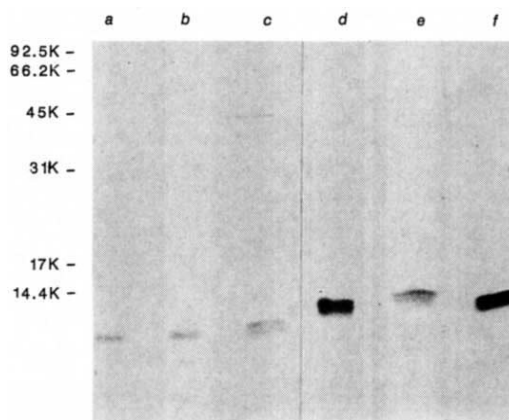
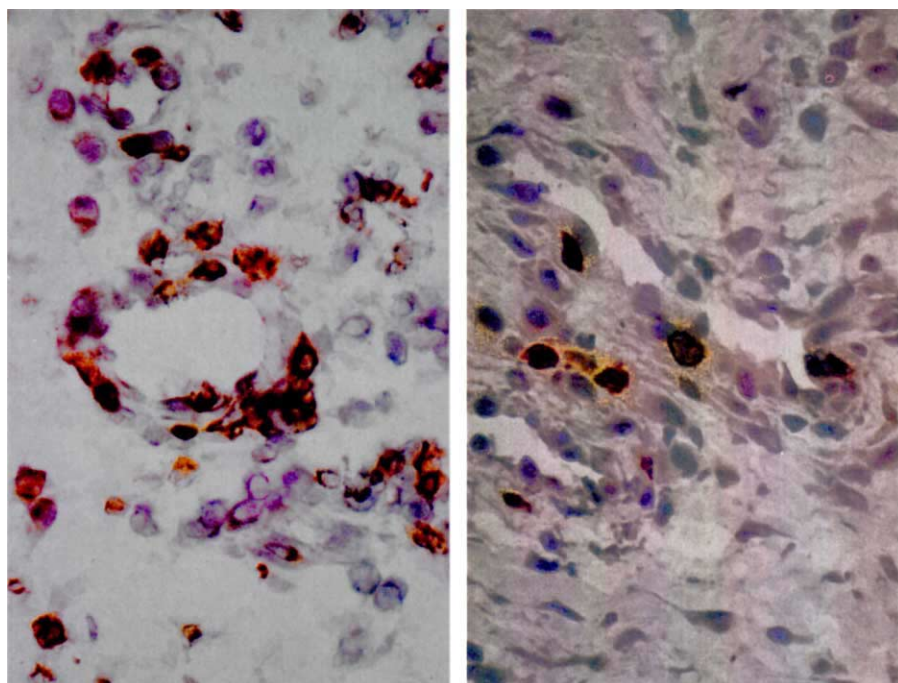


Fig. 4 Detection of MRP-8 and MRP-14 positive macrophages in histological sections of human rheumatoid polyarthritic synovialis. Many macrophages near blood vessels stain red for MRP-8 (left) or MRP-14 (right), other cells appear in blue.

Methods. Fixed cryosections (5 μ m) of human rheumatoid polyarthritic synovialis were fixed with acetone and stained by indirect immunoperoxidase¹⁴, using anti-MRP-8 and anti-MRP-14 sera. Mayer's hemalaun (blue) was used for counterstaining.



MRP-14 both contain a single cysteine residue, have no signal or membrane-anchor sequences, and lack consensus sequences for *N*-linked glycosylation. It can be predicted from the presence of stretches of amino acids that form Ca^{2+} -binding domains³, and shown also by $^{45}\text{Ca}^{2+}$ exchange on western blots (J. Lucas, personal communication), that MRP-8 and MRP-14 are Ca^{2+} -binding proteins. Both MRP-8 and MRP-14 contain two Ca^{2+} -binding domains which are quite distinct (Fig. 2). Recombinant MRP-8 and MRP-14 produced in *E. coli* comigrate with natural MRP-8 and MRP-14 (Fig. 1). The recombinant and natural proteins, whether crude or partially purified, individually or in combination, did not display or inhibit MIF activity, and did not react with the monoclonal antibody 1C5 used to isolate the natural proteins. Preliminary evidence suggests that natural MRP-8 and MRP-14 form a complex with a third component, carrying the 1C5-epitope, that is not eluted from the column at low pH with MRP-8 and MRP-14 (ref. 4).

Antisera were proven to be monospecific by Western blot analysis of recombinant MRP-8 and MRP-14 (Fig. 3, tracks *a* and *d*) and of the natural proteins (Fig. 3, tracks *b* and *e*), as well as macrophage lysates (tracks *c* and *f*). An immunohistochemical study⁵ has shown that the antisera stained cells of the myeloid lineage, granulocytes, monocytes and macrophages, exclusively. In peripheral and organ blood granulocytes and monocytes stained positive for MRP-8 and

MRP-14, whereas lymphocytes and platelets did not. In sections of normal human tissues neither protein was detectable in resident macrophages but in the acutely inflamed tissue found in gingivitis, psoriasis, neurodermitis and erythrodermia, MRP-14 was present in macrophages of the perivascular infiltrate in 15 out of 32 cases analysed, whereas MRP-8 was always absent. In contrast, in chronically inflamed tissue, particularly in primary chronic polyarthritis (11 cases so far), macrophages staining positive for MRP-8 are found in addition to MRP-14-positive macrophages (Fig. 4).

These data indicate that MRP-8 and MRP-14 are involved in chronic inflammatory reactions. The presence of MRP-8-positive macrophages in inflammatory lesions is in fact indicative of the chronic nature of the inflammatory process. Ca^{2+} -binding proteins have been implicated in macrophage activation^{6,7}, and the expression of these two proteins, and their interaction with each other and with other proteins⁴ could be important in the regulation of macrophage differentiation and to our understanding of the mechanisms underlying inflammatory disease.

We thank R. von Essen and H. Topoll for human biopsies and R. Knecht, H. Rink and W. Wiesendanger for their valuable contributions to this work. We thank J. Lucas for his permission to cite unpublished data. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00286.

Received 27 July; accepted 15 September 1987.

- Burmeister, G., Tarcsay, L. & Sorg, C. *Immunobiology* **171**, 461-474 (1986).
- Dorin, J. R. *et al.* *Nature* **326**, 614-617 (1987).
- Kretsinger, R. H. *CRC Crit. Rev. Biochem.* **8**, 119-174 (1980).
- Sorg, C. & Odink, K. in *Molecular Basis of Lymphokine Action* (eds Webb, D. R., Pierce, C. & Cohen, S.) (Humana, New Jersey, 1987).
- Zwadlo, G., Brüggens, J., Gerhards, G., Schlegel, R. & Sorg, C. *Clin. exp. Immun.* (in the press).
- Wright, B., Zeidman, I., Greig, R. & Poste, G. *Cell. Immun.* **95**, 46-53 (1985).
- Hartung, H. P. *FEBS Lett.* **160**, 209-212 (1983).
- Eur. Pat. Appl.* 162812 (1985).
- Rink, H., Liersch, M., Sieber, P. & Meyer, F. *Nucleic Acids Res.* **12**, 6369-6385 (1984).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schütz, G. *Nucleic Acids Res.* **9**, 2251-2266 (1981).
- Sanger, F., Nicklen, S. & Coulson, A. R., *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Feige, U., Overwien, B. & Sorg, C. *J. Immunol. Meth.* **54**, 309-315 (1982).
- Zwadlo, G., Schlegel, R. & Sorg, C. *J. Immun.* **137**, 512-519 (1986).

Vitamin K-dependent blood coagulation proteins form hetero-dimers

K. Harlos, S. K. Holland, C. W. G. Boys, A. I. Burgess*, M. P. Esnouf* & C. C. F. Blake

Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK
* Nuffield Department of Clinical Biochemistry, Radcliffe Infirmary, Oxford OX2 6HE, UK

The later stages of the blood coagulation cascade are characterized by the presence of vitamin K-dependent proteins and their involvement in membrane-bound, multi-protein converting complexes with an essential requirement for calcium ions. Specific interactions between zymogens and activating enzymes have not yet been identified. Here we describe a crystallographic study of prothrombin fragment 1 (residues 1-156 of prothrombin) which indicates that vitamin K-dependent coagulation proteins have specific association sites that allow them to form hetero-dimers. The calcium-induced formation of a hetero-dimer between fragment 1 and factor X is demonstrated by cross-linking. Such hetero-dimers of vitamin K-dependent proteins could be significant in the coagulation system.

The zymogens of blood coagulation and fibrinolysis are composed of modules of discrete structure and function^{1,2}. The number of module types is limited and different zymogens tend to have modules of the same type but often in different combinations. The vitamin K-dependent plasma proteins (prothrombin, factors X, IX and VII and proteins C, S and Z) contain an amino-terminal module of about fifty amino-acid residues with a high degree of homology (Fig. 1). This module contains all the γ -carboxyglutamic acid (Gla) residues³ which are formed by a vitamin K-dependent post-translational modification and are responsible for the high-affinity binding of Ca^{2+} ions⁴.

The crystal structure of prothrombin fragment 1 (refs 5 and 6) shows that the vitamin K-dependent region is divided into two structural parts: residues 1-35 are not represented by significant electron density and must be considered as spatially disordered, whereas residues 36-48 are well-ordered and form an α -helix. This structural division of the vitamin K-dependent region is reflected in the gene arrangement residues 1-38 and

Table 1 Possible functional hetero-dimers in blood coagulation system

Hetero-dimer	Cofactor
Factor X _a -prothrombin	Factor V _a
Factor IX _a -factor X	Factor VIII _a
Factor VII _a -factor X	Tissue factor III
Factor VII _a -factor IX	Tissue factor III
Protein C _a -protein S	Factor V _a and factor VIII _a

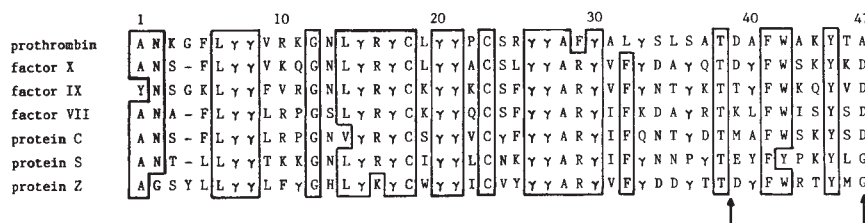
39-47 being encoded by different exons⁷⁻¹¹. The pattern of amino acid distribution (Fig. 1) is also divided in that residues 1-36 contain most γ -carboxyglutamic acid residues whereas the region containing residues 38-47 is characterized by the group of invariant hydrophobic aromatic-amino acid residues at positions 41, 42 and 45. The crystal structure of fragment 1 (ref. 5) shows that Phe 41, Trp 42 and Tyr 45 form a hydrophobic cluster on the outer edge of a helix. In the crystal these residues are not exposed to the solvent but interact around a crystallographic twofold axis with the equivalent region in a neighbouring molecule (Fig. 2a) to form a tightly linked dimer. The kringle of fragment 1 (residues 66-144) is not involved in this interaction.

The presumed function of the Gla region in binding calcium was investigated by soaking the crystals in 10 mM SrCl_2 solution (strontium is functionally equivalent to calcium¹², but easier to detect in electron density maps). The difference Fourier map at 6-Å resolution (Fig. 2b) shows one major and several minor peaks. The major peak is located on the crystallographic two fold axis to the amino-terminal side of residue 36 and is therefore associated with the apparently disordered Gla-containing region. This suggests that either this region is sufficiently ordered to bind added heavy metal ions, or that the addition of the metal ions increases the order in this region. Because the strontium density is situated on the twofold axis, the most probable form of binding will be in the form of Sr^{2+} bridges between equivalent Gla residues on different molecules. The protein elements closest to the Sr^{2+} peak are likely to be residues 25-33, which include Gla residues at positions 33, 30, 27 and 26 as potential ligands (Fig. 2b). The binding of Sr^{2+} ions thus extends the dimer interface and thereby strengthens the interaction between the two monomers.

The structural observation of a hydrophobic site between residues 40 and 45 is supported by biochemical studies¹³⁻¹⁶ and there is evidence that the fragment 1 dimer in the crystal is similar to the prothrombin dimer found in solution^{12,17}. The residues involved in this dimerization are highly conserved in all vitamin K-dependent blood coagulation proteins (Fig. 1), and they would therefore be expected to form dimers in the presence of calcium ions and to be able to associate with any other protein of this group to form a hetero-dimer.

To confirm the presence of hetero-dimers in solution, mixtures of fragment 1 and factor X were cross-linked (Fig. 3). The cleavable cross-linking reagent dithiobis(succinimidylpropionate) (DSP) has previously been used to investigate the association of fragment 1 (ref. 18) and prothrombin^{12,17}, being presumed to cross-link at Lys 44 (ref. 17). In the case of a fragment 1-factor X hetero-dimer, DSP would thus be expected to cross-link Lys 44 of fragment 1 to Lys 44 of factor X (prothrombin

Fig. 1 Sequence comparison of the amino-terminal region of vitamin-K dependent plasma proteins: bovine prothrombin³, bovine factor X (ref. 27), bovine factor IX (ref. 28), human factor VII (ref. 29), bovine protein C (ref. 30), bovine protein S (ref. 31) and bovine protein Z (ref. 32). Arrows mark the intron positions. γ , γ -carboxyglutamic acid (Gla) residues.



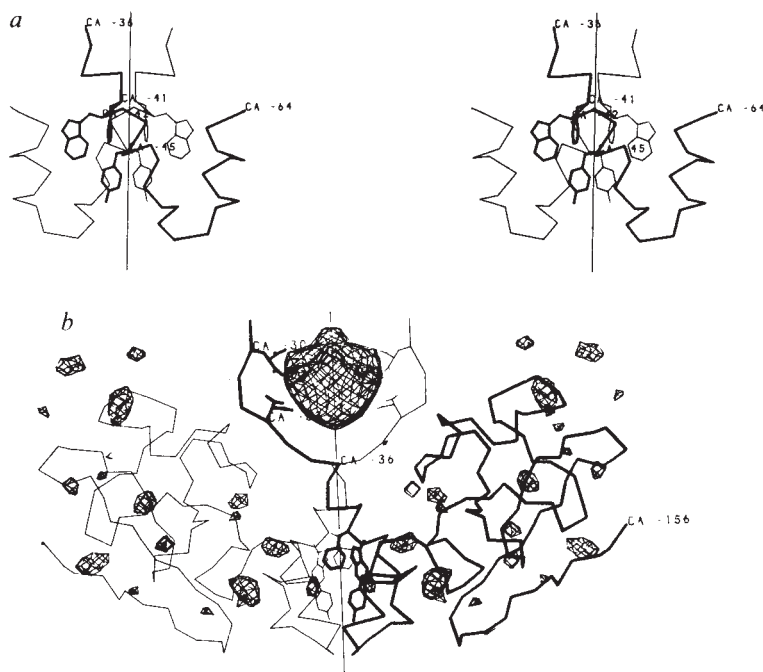


Fig. 2 *a*, Stereo view of the hydrophobic interaction of prothrombin fragment 1 molecules. Residues Phe 41, Trp 42 and Tyr 45 interact around a crystallographic twofold axis with the corresponding residues of a symmetry-related molecule to form a dimer in the crystal. These residues are highly conserved and are also present in other blood coagulation proteins (see Fig. 1). Crystals of bovine prothrombin fragment 1 were grown in the absence of divalent cations and the structure solved by isomorphous replacement and solvent flattening at 3.8-Å resolution⁵. The region shown is from residue 36–64. *b*, Difference Fourier map showing the binding of strontium ions to fragment 1. The largest feature of electron density is centred on the crystallographic twofold axis close to the fragment 1 molecules which form the dimer in the crystal. The difference Fourier was calculated using data to 6-Å resolution (815 reflections) of crystals which had been soaked in 10 mM SrCl₂ for 24 h. The position of residues 36–156 was taken from the native 3.8-Å map. Residues 29–35 were added to the structure to illustrate the possible interaction of Glu residues 30 and 33 with the strontium density.

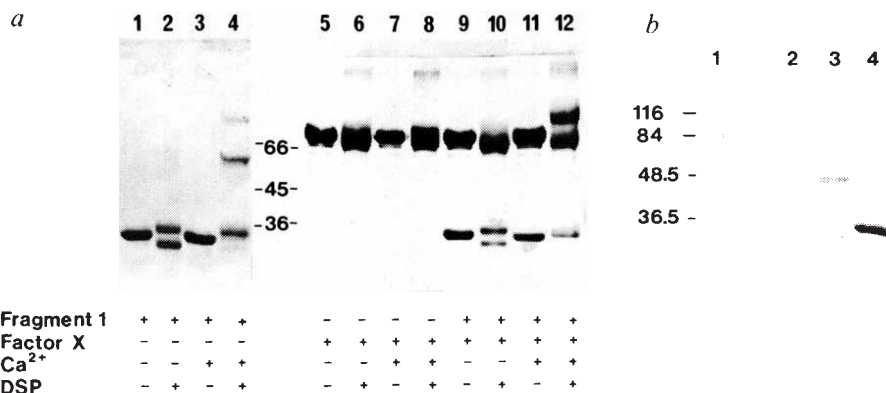
numbering). Control experiments with fragment 1 alone show that the fragment 1 dimer and tetramer can be detected. A new intense band with an apparent relative molecular mass (M_r) of 84,000 (84K) is formed only in the presence of Ca²⁺ and only if both fragment 1 and factor X are present (Fig. 3*a*, lane 12). This band can be isolated by high-performance size-exclusion chromatography (Fig. 3*b*, lane 1) and then analysed by treatment with 2-mercaptoethanol, which reduces the disulphide bond of DSP, thus breaking the cross-link. Factor X has a heavy and a light chain, joined by a disulphide bond¹⁹, which can be dissociated after reduction with mercaptoethanol. Reduction of the fragment 1-factor X hetero-dimer complex produces three bands of comparable intensity (Fig. 3*b*, lane 2), as would be predicted from the model. Control experiments (Fig. 3*b*, lanes 3 and 4) demonstrate that these bands correspond to fragment 1 and the heavy and light chains of factor X, which is a clear demonstration

that the 84K band is a one-to-one complex between fragment 1 and factor X.

The electrophoresis results give a preview of the equilibrium of the monomer-dimer and the monomer-hetero-dimer interactions under these conditions. During the dimerization of fragment 1 (Fig. 3*a*, lane 4) some 40% of all monomers associate. Factor X (lane 8) does not form a dimer to any significant extent. In an equimolar mixture of fragment 1 and factor X (lane 12), a large proportion is present in the monomeric form. The concentration of the fragment 1 dimer is, however, greatly reduced and is only just visible on the original gels. This indicates that the hetero-dimer is formed much more readily than either homo-dimer.

The hetero-dimer interface we describe has some interesting features. The strontium-binding experiment suggests that the Glu-containing region becomes ordered and therefore capable

Fig. 3 Electrophoretic analysis of the fragment 1-factor X hetero-dimer. *a*, the results of cross-linking experiments of mixtures of bovine prothrombin fragment 1 and bovine factor X in the presence or absence of Ca²⁺. The individual components of each lane are given at the bottom. Lanes contain equal amounts of fragment 1 and/or factor X for easy comparison. *b*, the isolated hetero-dimer complex before (lane 1) and after (lane 2) treatment with 2-mercaptoethanol to cleave the cross-linkage. Lanes 3 and 4 are control experiments containing in lane 3: factor X, Ca²⁺, DSP and in lane 4: fragment 1, Ca²⁺, DSP under reducing conditions. The M_r s of marker proteins are given on the left (K).



Methods. Bovine prothrombin fragment 1 and bovine factor X were isolated as described^{33,34}, and cross-linked with DSP (ref. 12). Approximately equimolar amounts of prothrombin fragment 1 and factor X were mixed and incubated in the presence or absence of Ca²⁺ for 30 min at room temperature, then DSP was added for the cross-linking reaction at 4°C for 45 min. The cross-linking was terminated with 0.1 M lysine. The final concentrations were 0.9 mg ml⁻¹ fragment 1, 1.8 mg ml⁻¹ factor X, 2.5 mM CaCl₂, 50 mM triethanolamine/HCl buffer, pH 7.5, 1 mM DSP and 5 mM lysine. The observed band splitting of fragment 1 (Fig. 3*a*, lane 2) is not a cleavage of fragment 1, but is probably due to some intermolecular crosslinkage of DSP. The band with apparent M_r 84K was present only in lane 12, and was isolated by high-performance size-exclusion chromatography using a TSK-2000 column with a flow rate of 0.4 ml min⁻¹ and a 0.15 M NaCl, 50 mM Tris/HCl buffer at pH 7.5. The eluted protein fraction was concentrated by freeze-drying. Large amounts of NaCl from the buffer solution remaining in the sample were removed by microdialysis. Electrophoresis was carried out with an LKB midiget system on SDS/10% polyacrylamide gels which were stained with Coomassie blue.

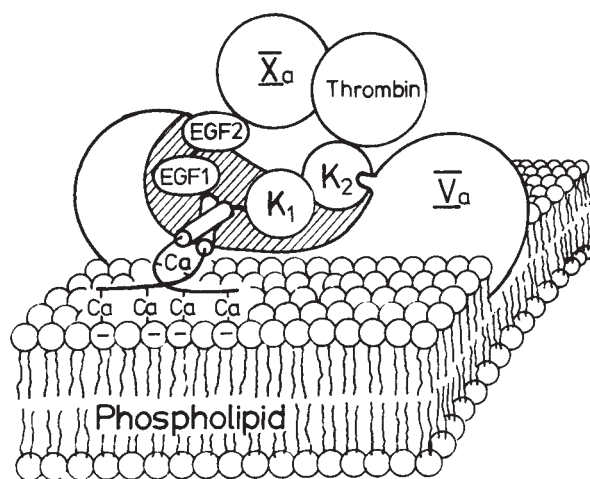


Fig. 4 Schematic representation of the prothrombinase complex, showing the interactions between prothrombin, factor X_a and factor V_a on the phospholipid surface. EGF1 and EGF2 are the two epidermal growth factor units of factor X_a , K1 and K2 are the kringle parts of prothrombin. The model was adapted from Nesheim *et al.*²³ to illustrate how prothrombin and factor X_a interact through their calcium-binding region to form a hetero-dimer complex. The proposed mechanism includes a hydrophobic interaction on the outer edge of a helix and a calcium-bridging mechanism between Gla residues. The Gla residues towards the N-terminal end attach the hetero-dimer complex to the phospholipid surface.

of dimer formation only in the presence of calcium. The hydrophobic site may not be present in solution and could only be formed during the Ca^{2+} -induced dimerization. The presence of a hydrophobic site ensures that calcium ions bridge equivalent Gla-residues in a parallel rather than in an antiparallel^{20,21} direction, thereby organizing the dimerization.

In Table 1 we list a number of steps of the blood coagulation cascade which could include hetero-dimer interactions. The equimolar complex of activated protein C and protein S (ref. 22) directly confirms our hetero-dimer model. Interestingly, this complex seems to exist in the absence of factor V_a or $VIII_a$. Our hetero-dimer model does not distinguish between different vitamin K-dependent plasma proteins and therefore permits all possible homo-dimer and hetero-dimer combinations. Under physiological conditions, however, only a few hetero-dimer complexes are functional in promoting blood coagulation, suggesting that additional factors influence the formation of the hetero-dimer.

One way to achieve functional specificity is through cofactor proteins with specific binding sites for the individual zymogen and activating enzyme. In the prothrombinase complex^{23,24} (see Fig. 4) the heavy chain of factor V_a binds to kringle 2 of prothrombin^{25,26}. But even in the absence of cofactors some molecule pairs are formed more readily than others: as shown here, fragment 1 and factor X have some additional, inherent specificity favouring the formation of the hetero-dimer over the homo-dimers.

We thank Dr D. I. Stuart for his help with the data processing and the MRC for financial support. K.H. was supported by the MRC, the Deutsche Forschungsgemeinschaft and the EPA Cephalosporin Fund. M.P.E. and C.C.F.B. are members of the Oxford Enzyme Group which is supported by the SERC.

Received 6 July; accepted 28 September 1987.

1. Patthy, L. *Cell* **41**, 657-663 (1985).
2. Doolittle, R. F. *TIBS* **10**, 233-237 (1985).
3. Magnusson, S., Peterson, T. E., Sottrup-Jensen, L. & Claeys, H. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 123-149 (Cold Spring Harb., New York, 1975).
4. Nelsestuen, G. L. *J. biol. Chem.* **251**, 5648-5656 (1976).

5. Harlos, K., Boys, C. W. G., Holland, S. K., Esnouf, M. P. & Blake, C. C. F. *FEBS Lett.* (in the press).
6. Park, C. H. & Tulinsky, A. *Biochemistry* **25**, 3977-3982 (1986).
7. MacGillivray, R. T. A. & Irwin, D. M. *Haemostasis* **16**, 227-238 (1986).
8. Leytus, S. P., Foster, D. C., Kurachi, K. & Davie, E. W. *Biochemistry* **25**, 5098-5102 (1986).
9. Anson, D. S. *et al. EMBO J.* **3**, 1053-1060 (1984).
10. Foster, D. C., Yoshitake, S. & Davie, E. W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4673-4677 (1985).
11. O'Hara, P. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 5158-5162 (1987).
12. Tarvers, R. C., Noyes, C. M., Roberts, H. R. & Lundblad, R. L. *J. biol. Chem.* **257**, 10708-10714 (1982).
13. Tarvers, R. C. *Archs Biochem. Biophys.* **241**, 639-648 (1985).
14. Brenckle, G. M., Peng, C. W. & Jackson, C. M. in *Vitamin K Metabolism and Vitamin K-dependent Proteins* (ed. Suttie, J. W.) 54-57 (University Park, Baltimore, 1980).
15. Bernarous, R., Gacon, G., Rabiet, M. & Labie, D. in *Vitamin K Metabolism and Vitamin K-dependent Proteins* (ed. Suttie, J. W.) 49-53 (University Park, Baltimore, 1980).
16. Sarasua, M. M. *et al. Biochim. biophys. Acta* **742**, 257-264 (1983).
17. Tarvers, R. C., Noyes, C. M., Tarvers, J. K. & Lundblad, R. L. *J. biol. Chem.* **261**, 4855-4859 (1986).
18. Tarvers, R. C., Roberts, H. R. & Lundblad, R. L. *J. biol. Chem.* **259**, 1944-1950 (1984).
19. DiScipio, R. G., Hermodson, M. A. & Davie, E. W. *Biochemistry* **16**, 5253-5260 (1977).
20. Jackson, C. M., Peng, C., Brenckle, G. M., Jonas, A. & Stenflo, J. *J. biol. Chem.* **254**, 5020-5026 (1979).
21. Mann, K. G., Prendergast, F. G. & Bloom, J. W. in *The Regulation of Coagulation* (eds Mann, K. G. & Fletcher, B. T.) 3-9 (Elsevier, New York, 1980).
22. Walker, J. F. *J. biol. Chem.* **256**, 11128-11131 (1981).
23. Nesheim, M. E. *et al. in The Regulation of Coagulation* (eds Mann, K. G. & Fletcher, B. T.) 145-155 (Elsevier, New York, 1980).
24. Mann, K. G. & Esnouf, M. P. in *Haemostasis and Thrombosis* (eds Bowie, E. J. W. & Sharp, A. A.) 148-172 (Butterworths, London, 1985).
25. Esmon, C. T. & Jackson, C. M. *J. biol. Chem.* **249**, 7791-7797 (1974).
26. Guinto, E. R. & Esmon, C. T. *J. biol. Chem.* **259**, 13986-13992 (1984).
27. Enfield, D. L. *et al. Biochemistry* **19**, 659-667 (1980).
28. Katayama, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4990-4994 (1979).
29. Hagen, F. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2412-2416 (1986).
30. Fernlund, P., Stenflo, J. & Tufvesson, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5889-5892 (1978).
31. Dahlbäck, B., Lundwall, A. & Stenflo, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4199-4203 (1986).
32. Højrup, P., Jensen, M. S. & Petersen, T. E. *FEBS Lett.* **184**, 333-338 (1985).
33. Aschaffenburg, R., Blake, C. C. F., Burridge, J. M. & Esnouf, M. P. *J. molec. Biol.* **114**, 575-579 (1977).
34. Esnouf, M. P., Lloyd, P. H. & Jesty, J. *Biochem. J.* **131**, 781-789 (1973).

Calculation of electrostatic potentials in an enzyme active site

Michael K. Gilson & Barry H. Honig

Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, New York 10032, USA

To be able to calculate the contributions of individual amino acids to the electrostatic field of a protein would be of considerable value in designing proteins of enhanced or altered function and stability. Recent studies on the serine protease subtilisin provide direct measurements of the electrostatic potential in the active site of the enzyme produced by two charged amino acids^{1,2}. We have used these results to test a recently developed method for the calculation of electrostatic interactions between two specific sites on a protein³. The extent of agreement between the theoretical and experimental results suggests that the continuum solvent model used in the calculations reproduces the essential features of the interaction.

Continuum models for the interactions of charges in macromolecules are founded in classical electrostatic theory, which treats materials as homogeneous dielectric media that may be polarized by electrical charges. Systems composed of charges and dielectric media are described by Poisson's equation, $\nabla \cdot (\epsilon(\mathbf{r}) \nabla \phi(\mathbf{r})) = -4\pi\rho(\mathbf{r})$, where $\epsilon(\mathbf{r})$ is the dielectric constant as a function of position, $\phi(\mathbf{r})$ is the electrostatic potential and $\rho(\mathbf{r})$ is the charge density. Proteins have low dielectric constants (about 2-3), because reorientation of their dipolar groups is severely restricted⁴. Because water has a very high dielectric constant (~ 80), the protein-water interface constitutes a boundary between two dielectric media. The atomic coordinates of a protein, obtained by crystallography or by model-building, define the shape of this dielectric boundary, as well as the locations of ionized or partially charged atoms. Because both the charge distribution and dielectric constant are defined everywhere in space, it is possible, at least in principle,

to solve Poisson's equation for a protein and obtain the electrostatic field as a function of position. A full solution yields charge-charge interaction energies, as well as charge-solvent interaction energies which account for the free energy of transferring charged atoms between solvent and the macromolecular interior^{5,6}. If the water contains an electrolyte, as is normally the case in systems of physiological interest, the Poisson-Boltzmann equation⁷ should be used, because this accounts for the tendency of mobile ions to redistribute themselves according to the electrical potential in solution. In such calculations, the ionic strength may be set to any desired value in the solvent region, but should be set to zero in the interior of the molecule, as the protein excludes solvent.

The implementation of numerical methods on high-speed computers makes it possible to solve the Poisson-Boltzmann equation for molecules of arbitrary shape and charge distribution^{3,8,9}. Such methods have been applied to proteins in studies of substrate diffusion in the electric field of Cu, Zn-superoxide dismutase^{3,10}, the interactions of Klenow fragment of DNA polymerase with B-DNA (ref. 11), and the effect of a specific charge on the redox properties of cytochrome *c*₅₅₁ (ref. 12).

Fersht and co-workers^{1,2} have recently performed a series of measurements on the enzyme subtilisin BPN' which serve as excellent tests for models of electrostatic interactions. Site-directed mutagenesis was used to substitute neutral residues for charged ones, and the resulting *pK* changes in the active site His (His 64) were measured at six ionic strengths, by analysis of the kinetics of the enzyme. As pointed out by Fersht and co-workers, these shifts should be proportional to the changes in electrostatic potential at the His on going from wild type to mutant. In two series of experiments, either Asp 99 or Glu 156, each roughly 13 Å from the catalytically essential His 64, was replaced with Ser. We have attempted to account for the observed *pK* shifts by means of a continuum electrostatic model, solved using a modified form⁵ of the algorithm of ref. 3. The primary modification is the introduction of a Stern layer⁷, which prevents van der Waals overlap between dissolved ions and protein atoms at the interface between protein and solvent. (See also Table 1 legend.)

Overall, the agreement between calculated and measured values is good (Table 1). The calculated Glu 156 potentials fall close to the experimental values and do not show a systematic error. The calculated Asp 99 potentials are somewhat lower than the measured potentials at the lower ionic strengths. Regarding experimental error, it should be noted that the Asp 99 measurements at 0.10 M ionic strength were performed under three different experimental conditions which should yield the same His 64 *pK* shifts. The observed shifts are -0.29 (0.03), -0.23 (0.03) and -0.26 (0.02) (the numbers in parentheses being the reported standard deviations). Taken together, these give -0.26 (0.06), suggesting that the experimental uncertainty is about 0.06.

The calculated *pK* shifts are somewhat insensitive to the dielectric constant assumed for the protein. For example, changing the protein dielectric constant from 2 to 10 changes the His 64 *pK* shift due to the Asp 99 mutation at 0.10 M ionic strength from -0.18 to -0.15. Thus any inhomogeneity of the protein dielectric constant should have little effect on the interactions. The basis for this insensitivity to protein dielectric constant is that, for groups near the protein surface, most of the interaction is through water rather than through protein. This may be understood intuitively from classical electrostatics where it is known that lines of electric displacement (**D**) tend to concentrate in high- rather than low-dielectric regions. Because the electric field is given by **D**/ ϵ , the large **D** flux in the high-dielectric solvent will be heavily screened. This phenomenon can be seen in model calculations, where the effective dielectric constant for interactions between groups on the protein surface will, in general, be high even though the actual dielectric constant of the protein is low⁵.

Table 1 *pK* shifts for mutant subtilisins as a function of ionic strength

Ionic strength		<i>pK</i> shift	
		Asp 99 → Ser	Glu 156 → Ser
0.500	Calculated	0.10	0.19
	Experimental	0.09	NA
0.100	Calculated	0.18	0.27
	Experimental	0.23, 0.26, 0.29	0.25
0.025	Calculated	0.25	0.34
	Experimental	0.36	0.41
0.010	Calculated	0.29	0.37
	Experimental	0.42	0.42
0.005	Calculated	0.31	0.39
	Experimental	0.38	—*
0.001	Calculated	0.34	0.42
	Experimental	NA	0.39†

Calculated and measured (ref. 2) *pK* shifts of His 64 on going from wild type (D99, E156) to either S99 or S156 mutants, as a function of ionic strength. (Shifts multiplied by -1 for clarity.) Experimental values for Asp 99 at 0.1 M ionic strength are as follows: middle entry, standard conditions, as for other measurements; first entry measured using a different substrate; third entry measured in the presence of 0.1 mM EGTA. NA, not available. The Linearized Poisson-Boltzmann equation solved assuming a unit charge on His 64 N_{e2} , yielding potential throughout the protein. The potential at the carboxylates of Asp 99 and Glu 156 must be equal to the potential that would exist at N_{e2} of His 64 if the respective carboxylates had been charged. The His 64 *pK* shift due to mutating each carboxylate is $-(0.5\phi_1 + 0.5\phi_2)/1.4$, where ϕ_1 and ϕ_2 are the potentials at the two oxygens (kcal per mol per electron change). Protein interior dielectric constant, 2; solvent dielectric constant, 78.5. Thickness of the Stern layer, 2 Å. Protein surface calculated using united atom Van der Waals radii¹⁷. Protein interior includes bound waters of crystal structure. Finite difference algorithm is that of Klapper *et al.*³, except that focusing was used to improve boundary conditions⁸; solutions should be correct within about 10% (ref. 8). Brookhaven Protein Data Bank provided atomic coordinates of subtilisin (1SBT). This structure has Ala 99 instead of the correct Asp 99, so Asp side chain was built in with the program PAKGGRAF¹⁸, keeping the main-chain atoms fixed, and avoiding steric overlaps. Resulting atomic coordinates: C γ : 7.25, 21.37, 25.57; O δ_1 : 6.86, 20.78, 24.57; O δ_2 : 6.67, 22.64, 25.88.

*Experimental value questionable (A. Fersht, personal communication).

†Data from A. Fersht.

In the calculations reported in Table 1, waters seen in the crystal structure, some of which lie in the active-site cleft, are treated as part of the low-dielectric protein interior. When the calculations are repeated with these waters treated as part of the high-dielectric solvent, the agreement with experiment is not as good, particularly for Glu 156. For example, at low ionic strength (0.001 M for Glu 156, 0.005 M for Asp 99) the calculated His 64 *pK* shifts due to Glu 156 and Asp 99 change by +0.15 and -0.01 respectively, and at 0.1 M, the shifts change by +0.11 and -0.02 respectively. It is initially surprising that decreasing the quantity of low-dielectric solvent in the active site should actually increase the calculated potential produced by Glu 156 at His 64. However, at least one of the bound waters lies between Glu 156 and His 64, and it is known that a low-dielectric body between two charges in high-dielectric solvent can actually weaken their interaction^{5,13}. This results from the fact, mentioned above, that electric field lines avoid regions of low dielectric constant. A similar effect has recently been noted in the case of DNA as shown by Zimm and co-workers¹⁴.

There are a number of simpler models for electrostatic interactions, and it is worth determining whether these reproduce the experimental data. The simplest is Coulomb's law, which is a solution of the Poisson-Boltzmann equation for the case where no dielectric boundaries are present and the ionic strength is zero. This is often used in molecular mechanics calculations along with a distance dependent dielectric constant, $\epsilon = R$, which accounts approximately for solvent screening of electrostatic interactions. Coulomb's law with $\epsilon = R$ markedly overestimates

the pK shifts, yielding -1.21 and -1.53 for the shifts induced by Asp 99 and Glu 156, respectively. This result agrees with a recent observation that this model appears to overestimate electrostatic interactions in the protein crambin¹⁵. A more complex screening formula suggested by Warshel and co-workers¹⁶ yields the overestimates -0.51 and -0.61 for the same shifts. Neither $\epsilon = R$ nor the formula suggested in ref. 16 gives results which depend on ionic strength. A simple model which does account for ionic strength is the special case solution to the linearized Poisson-Boltzmann equation incorporated into Debye-Huckel theory⁷, which yields the electrostatic potential as a function of distance from a spherical ion in an electrolyte: $\phi(r) = (q/\epsilon r) \exp(-\kappa r) \{ \exp(\kappa a)/(1 + \kappa a) \}$. Here r is the radial distance from an ion of radius a and charge q , which is immersed in a solvent having dielectric constant ϵ and ionic strength determined by the Debye-Huckel parameter κ . (The other symbols have their usual meanings.) When applied to subtilisin, this model neglects the facts that the protein does not contain any electrolyte and that it has a low dielectric constant. As a consequence, it underestimates the electrostatic interactions, yielding pK shifts of about 0.12 for both mutations at 0.01 M ionic strength.

In summary, the linearized Poisson-Boltzmann model used

here is found to account for the magnitude and ionic strength dependence of the electrostatic interactions observed by Fersht and co-workers in subtilisin. The success of classical concepts as intuitive guides and as a means of calculating experimental observables suggests that they will be useful both in developing insight into the role of electrostatics in protein structure and function, and in protein design.

Received 1 June; accepted 17 September 1987.

1. Thomas, P. G., Russell, A. J. & Fersht, A. R. *Nature* **318**, 375-376 (1985).
2. Russell, A. J., Thomas, P. G. & Fersht, A. R. *J. molec. Biol.* **193**, 803-813 (1987).
3. Klapper, I., Hagstrom, R., Fine, R., Sharp, K. & Honig, B. *Proteins* **1**, 47-59 (1986).
4. Gilson, M. K. & Honig, B. *Biopolymers* **25**, 2097-2119 (1986).
5. Gilson, M. K., Rashin, A., Fine, R. & Honig, B. *J. molec. Biol.* **183**, 503-516 (1985).
6. Rashin, A. A. & Honig, B. *J. phys. Chem.* **89**, 5588-5593 (1985).
7. Bockris, O'M. & Reddy, A. K. N. *Modern Electrochemistry* (Plenum, New York, 1973).
8. Gilson, M. K., Sharp, K. & Honig, B. H. *J. comput. Chem.* (in the press).
9. Warwicker, J. & Watson, H. C. *J. molec. Biol.* **157**, 671-679 (1982).
10. Sharp, K., Fine, R. & Honig, B. *Science* **236**, 1460-1463 (1987).
11. Warwicker, J., Ollis, D., Richards, F. M. & Steitz, T. A. *J. molec. Biol.* **186**, 645-649 (1985).
12. Rogers, N. K., Moore, G. R. & Sternberg, M. J. E. *J. molec. Biol.* **182**, 613-616 (1985).
13. Finkel'shtein, A. V. *Molek. Biol.* **11**, 627-634 (1978).
14. Troll, M., Roitman, D., Conrad, J. & Zimm, B. H. *Macromolecules* **19**, 1186-1194 (1986).
15. Whitlow, M. & Teeter, M. M. *J. am. chem. Soc.* **108**, 7163-7172 (1986).
16. Warshel, A., Russell, S. T. & Chung, A. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4585-4789 (1984).
17. McCammon, J. A., Wolynes, P. G. & Karplus, M. *Biochemistry* **18**, 927-942 (1979).
18. Katz, L. & Levinthal, C. A. *Rev. Biophys. Bioengng.* **1**, 465-504 (1972).

Prediction of electrostatic effects of engineering of protein charges

Michael J. E. Sternberg*, Fiona R. F. Hayes*,
Alan J. Russell†, Paul G. Thomas† & Alan R. Fersht†

* Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK

† Department of Chemistry, Imperial College of Science and Technology, London SW7 2AY, UK

Accurate prediction of electrostatic effects on catalytic activity is an essential component of protein design¹⁻⁴. Site-directed mutagenesis of charged groups in subtilisin of *Bacillus amyloliquefaciens* has provided experimental measurements of electrostatic interactions which may be used to test such theoretical methods. The pK_a of the histidine of the active site has been perturbed by $+0.08$ to -1.0 units by modifying one or two residues⁵⁻⁷. Electrostatic effects in proteins can be modelled by the algorithm of Warwicker and Watson, which uses classical electrostatics and considers both the charge position and the shape of the molecule⁸⁻¹⁰. Here we report that the algorithm can model several pK_a shifts in subtilisin to fair accuracy.

Ab initio modelling of electrostatic effects in proteins requires consideration of the atomic polarizabilities⁴ of the heterogeneous protein and the solvent, including both water and counterions. The size of the system makes approaches of this sort difficult and so simplified methods are necessary, such as the use of a macroscopic dielectric (Fig. 1a). Simple models use uniform or distance-dependent dielectrics^{11,12}. These models ignore the importance of the exact location of the charges as there is far greater solvent screening of surface charges than buried charges. However, these features are incorporated into the Warwicker-Watson algorithm, in which the protein/solvent is divided into cubes, typically of dimension 1 Å and, depending on location, a dielectric (D) is assigned to each cube which corresponds to the protein ($D=3.5$), the solvent ($D=80$) or takes an intermediate value for the boundary (Fig. 1b). Each charge is spread to the eight corners of the cube. A finite difference solution to Poisson's equation is then obtained. The version of the algorithm was benchmarked¹⁰ by reproducing a

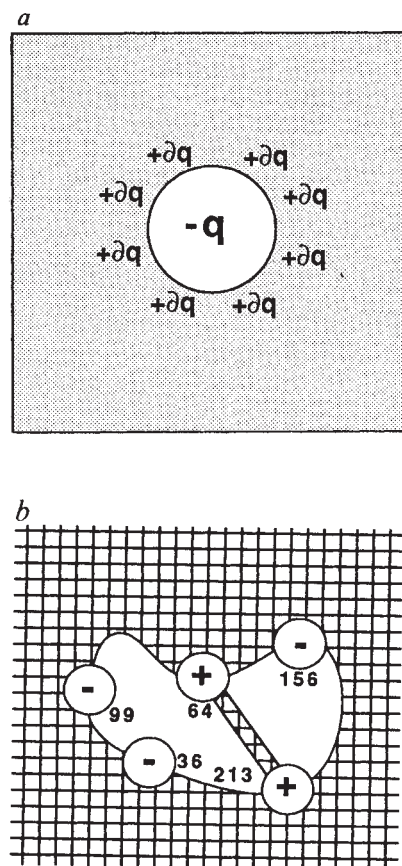


Fig. 1 a, The physics of the dielectric effect. A charge q in a continuous medium of dielectric D induces a surface polarization $\Sigma \delta q$ which partially neutralizes the original charge. The effective charge inside the dielectrics $-(q - \Sigma \delta q) = -q/D$. b, Calculation of electrostatic effects in mutant subtilisins. The Warwicker-Watson algorithm uses a grid and associates dielectrics and charges appropriately. The positions of the charged residues illustrate roughly whether the interaction is mainly through protein or solvent.

Table 1 The difference in pK_a between mutant and wild type in different buffers

Mutant	pK_a shift				Average
	$I = 0.001$ M imidazole	$I = 0.01$ M imidazole	$I = 0.005$ M P_i	$I = 0.01$ M P_i	
Asp → Ser 99	-0.40 ± 0.02	—	-0.38 ± 0.02	-0.42 ± 0.02	-0.40
Glu → Ser 156	-0.39 ± 0.02	—	-0.32 ± 0.02	-0.42 ± 0.02	-0.38
Lys → Thr 213	$+0.10 \pm 0.02$	—	$+0.08 \pm 0.02$	$+0.07 \pm 0.02$	+0.08
Asp → Gln 36	$-0.18 \pm 0.02^\dagger$	—	—	—	-0.18
Asp → Lys 99	-0.57 ± 0.02	-0.68 ± 0.02	-0.65 ± 0.03	-0.65 ± 0.04	-0.64
Glu → Lys 156	-0.62 ± 0.02	-0.63 ± 0.02	-0.65 ± 0.04	-0.61 ± 0.03	-0.63
Asp → Ser 99 and Glu → Ser 156	-0.65 ± 0.04	—	-0.56 ± 0.04	-0.66 ± 0.03	-0.62
Asp → Lys 99 and Glu → Lys 156	-1.00 ± 0.06	—	—	—	-1.00

The values of pK_a shift are pK_a (mutant) - pK_a (wild-type) in buffers of different ionic strength (I). Kinetic data was measured and the Lys → Thr 213 and Asp → Gln 36 mutants were prepared and purified as in refs 6 and 7. Primers: 5'-CCCCGTATG*TGTTTCCA (Thr 213) and 5'-TGAGAAGAC*TG*GATACCGC (Gln 36) were used to direct the mutations. †, Value for Asp → Gln 36 is the average of two measurements: -0.17 ± 0.02 and -0.19 ± 0.02 . All the substitutions except Asp → Lys 99 and Glu → Lys 156 are observed in different species of subtilisin.

Table 2 Prediction of pK_a shifts due to mutations

Mutant	Mean distance from charge to His 64 nitrogen atoms (Å)	Experimental		Calculated (W&W)*		Calculated ΔpK_a	
		ΔpK_a	D_{eff}	ΔpK_a	D_{eff}	$D = 55$	$D (M \& E)^\dagger$
Asp → Ser 99	12.6	-0.40	48	-0.31	62	-0.35	-0.38
Glu → Ser 156	14.4	-0.38	45	-0.35	48	-0.31	-0.30
Ser → Lys 99	15.0	(-0.25)	65	-0.25	65	-0.30	-0.28
Ser → Lys 156	16.5	(-0.25)	59	-0.30	49	-0.27	-0.24
Lys → Thr 213	17.6 (?)	+0.08	173	+0.19	73	+0.25	+0.21
Asp → Gln 36	15.1	-0.18	90	-0.16	101	-0.29	-0.27
Asp → Lys 99	(13.8)	-0.64	55	-0.56	63	-0.64	-0.64
Gly → Lys 156	(15.5)	-0.63	50	-0.65	48	-0.57	-0.52
Asp → Ser 99 and Glu → Ser 156	(13.5)	-0.63	57	-0.66	55	-0.66	-0.66
Asp → Lys 99 and Glu → Lys 156	(14.7)	-1.00	66	-1.21	55	-1.21	-1.14

Numbers in parentheses denote experimental values calculated from two or more other values rather than directly determined. Distances for mutants were obtained assuming the side chain to be fully extended. Values for the pK_a shifts are the mean values on the two histidine imidazole nitrogens. Distances are the average distance from the side chain nitrogen or oxygens to the two histidine imidazole nitrogens.

* In the Warwicker-Watson (W&W) calculation, the effects of the Asp, Glu or Lys charged side chains were evaluated and other side chains were taken to be uncharged. The first six mutants in the table were calculated directly by the Warwicker-Watson algorithm and the others obtained by addition.

† The values for the pK_a are given using a constant dielectric of 55 or using the distance-dependent equation of ref. 12 (M&E).

previous result that modelled the redox potential shift in cytochrome c_{551} due to the introduction of a charge on a buried propionate group as 87 mV (effective dielectric, D_{eff} , equals 20) compared with experimental values of 65 mV ($D_{eff} = 27$).

We examined whether the Warwicker-Watson algorithm can be used to model the effects of single and double site-specific mutations on the pK_a of the active site His 64 in *Bacillus amyloliquefaciens* subtilisin⁵⁻⁷ (Table 1). As the algorithm does not include the effect of counterions the theoretical results model experimental values at low ionic concentration. Table 1 shows there is little variation in the results over two to four determinations of the pK_a shift in phosphate and imidazole buffers over a range of 0.001 to 0.01 M.

Coordinates of the crystal structure of *B. amyloliquefaciens* subtilisin BPN' (M. James, personal communication) were used and additional unpublished distances in the mutant enzymes were provided by R. R. Bott. The structures of the mutant enzymes were modelled from the coordinates using the computer graphics program FRODO¹³ with side chains being built in the extended conformation. In the algorithm, the effect of charges on the calculated electrostatic potential is additive. Thus the

appropriate charge was placed on the side-chain atom and the electrostatic potential at His 64 evaluated. This potential can be related to the shift in pK_a by the Tanford-Roxby¹⁴ equation ($\Delta pK_a = e\Delta\Phi/2.303zKT$, where e is the electronic charge, $\Delta\Phi$ is the change in the electrostatic potential at the titrating site, z the charge at the titrating site (1 for His 64) and $T = 298$ K). In addition the shifts in pK_a can be expressed⁷ as an effective dielectric between the charges using $D_{eff} = 244/\Delta q r \Delta pK_a$, where r is the separation in Å and Δq is the change in charge.

Table 2 shows that for the single mutations, apart from Lys → Thr 213, the calculated pK_a shifts agree with the experimental values to within 0.1 of a unit, better than a factor of 2. Experimentally the pK_a shifts for the double mutants were found to be less than the sum of values for the component single mutations. As the modelling is based on addition of the individual residue changes, poorer agreement for these double mutants was obtained. Assuming additivity of electrostatic effects, the experimental results show a larger pK_a shift for Asp 99 (or Glu 156) to a neutral side chain than for the corresponding mutations of a neutral side chain to Lys 99 (or Lys 156). This is because the extended Lys side chain is farther from the active

site histidine than the shorter acidic side chains and this effect is correctly predicted in the modelling.

Classical electrostatic theory shows that in systems with at least two media of different dielectrics, it is possible for the effective dielectric for an interaction to be greater than any individual dielectric¹. An important feature of the Warwicker-Watson algorithm is that we find it can yield effective dielectrics greater than that of the solvent. Thus for the Asp→Gln 36 mutation, the experimental D_{eff} is 90 and this is modelled in the algorithm with a calculated D_{eff} of 101.

The worst prediction for a single mutation is Lys→Thr 213, where the experimental result is 0.08 compared to the calculated value of 0.23. However, this is a particularly flexible region of the molecule and in crystal structures of several mutant enzymes (R. R. Bott, personal communication) the $N\epsilon$ of Lys 213 is $19 \pm 4 \text{ \AA}$ from His 94. Modelling with a structure that has a separation of $23 \text{ \AA}$ rather than $17.6 \text{ \AA}$ and perhaps has the charge more exposed to solvent should yield a better agreement with experiment.

Different algorithms can be applied to model these effects. Independently, Gilson and Honig¹⁵ have shown that the pK_a shifts of the mutations Asp→Ser 99 and Glu→Ser 156 at a range of ionic concentrations may be satisfactorily calculated using a related procedure in the form of a linearized version of the Poisson-Boltzman equation and classical electrostatics. One simpler model, suitable for predicting subsequent mutations, is to use a uniform dielectric of 55, as this is the mean value of the experimental D_{eff} values that are less than 80 (that is, excluding Lys→Thr 213 and Asp→Gln 36). Another method is the distance-dependent dielectric of Mehler and Eichele¹², designed to model electrostatic effects from surface charges, which always uses a D_{eff} less than that of the solvent. Most mutations have an experimental $D_{\text{eff}} < 80$ and for these both simpler approaches yield calculated pK_a shifts that agree with experiment as well as the Warwicker-Watson algorithm (Table 2). However the limitations of these simpler methods are highlighted in modelling the pK_a shifts for the two mutations with an experimental $D_{\text{eff}} > 80$. The simpler methods give poorer agreement with experiment than the results obtained by the Warwicker-Watson algorithm. Although the use of any macroscopic dielectric to model microscopic electrostatic effects is suspect, we conclude that consideration of the protein charge position and shape together with the different dielectric responses of the protein and solvent^{8-10,15} provides an approach of modelling that yields good predictions to guide protein engineering.

We thank Dr M. N. G. James for unpublished coordinates of subtilisin, Dr R. R. Bott for distances in mutant subtilisin structures and Dr N. K. Rogers for the computer program of the Warwicker-Watson algorithm. Drs Gilson and Honig provided us with a copy of their paper before publication. This work was funded by the SERC Protein Engineering initiative. A.J.R. receives an SERC Biotechnology Directorate Studentship.

Received 30 June; accepted 17 September 1987.

1. Rogers, N. K. *Prog. Biophys. molec. Biol.* **48**, 37-66 (1986).
2. Tam, S. C. & Williams, R. J. P. *Struct. Bond.* **63**, 103-151 (1985).
3. Matthew, J. B. *et al. CRC crit. Rev. Biochem.* **18**, 91-197 (1985).
4. Warshel, A. & Russell, S. T. Q. *Rev. Biophys.* **17**, 283-422 (1984).
5. Thomas, P. G., Russell, A. J. & Fersht, A. R. *Nature* **318**, 375-376 (1985).
6. Russell, A. J., Thomas, P. G. & Fersht, A. R. *J. molec. Biol.* **193**, 803-813 (1987).
7. Russell, A. J. & Fersht, A. R. *Nature* **328**, 496-500 (1987).
8. Warwicker, J. & Watson, H. C. *J. molec. Biol.* **157**, 671-679 (1982).
9. Rogers, N. K. & Sternberg, M. J. E. *J. molec. Biol.* **174**, 527-542 (1984).
10. Rogers, N. K., Moore, G. R. & Sternberg, M. J. E. *J. molec. Biol.* **182**, 613-616 (1985).
11. Hopfinger, A. J. in *Conformational Properties of Macromolecules* 59-63 (Academic, New York, 1973).
12. Mehler, E. L. & Eichele, G. *Biochemistry* **23**, 3887-3891 (1984).
13. Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
14. Tanford, C. & Roxy, R. *Biochemistry* **11**, 2192-2198 (1972).
15. Gilson, M. K. & Honig, B. H. *Nature* **330**, 84-86 (1987).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Thermal analysis of superconductors

W.P. Brennan, M.P. DiVito, R.F. Culmo and C.J. Williams

Thermal analysis techniques have been used to study the stoichiometry of superconductors. As the critical temperatures of the materials rise, thermal analysis will open the door to further characterization.

NOTHING in recent years has captured the excitement and interest of physical scientists as has superconductivity¹⁻³. In research laboratories around the world, scientists are working on projects connected with the phenomenon. Many governments have recognized the strategic importance of superconductivity, and are providing substantial grants to industrial and government laboratories for the study of superconducting materials.

The race in superconductivity is to develop materials which superconduct at successively higher temperatures. The basic materials that are being studied currently may generally be purchased off-the-shelf. One entrepreneurial company even supplies kits and recipes for secondary schools at a very modest cost.

At the present time, much of the research centres on ceramics based upon rare- and alkaline-earth copper oxides. This paper will discuss briefly the chemical aspects of these reactions, the thermogravimetric analysis (TGA) characterization tests that may be employed, and a brief review of other potential applications of thermal analysis in the study of superconductors.

Chemical aspects

The revolution in the discovery of ceramic superconductors has changed the scientific community's understanding of the composition of superconducting materials. Previously, it was believed that metals were the best superconductors. The discovery that ceramics may superconduct was not only surprising, but was difficult to understand because it contradicted the normal conventions of the current understanding of superconductors. Further, the first breakthrough superconductor of an yttrium, barium, copper oxide composition ($\text{YBa}_2\text{Cu}_3\text{O}_7$, — the so called "1-2-3" composition) did not follow the conventions of mineral perovskites, which dictate that the material contain three oxygen atoms for every two metal atoms⁴. The first material had close to a 1:1 ratio of metal and oxygen atoms, and had two oxygen atoms less than the predicted structure. Perovskites also traditionally have a three-dimensional crystalline array, and this compound was shown to have a layered structure.

The difference in oxygen content, and the influence of oxygen on the superconducting properties of ceramic materials, have been the subjects of a great amount of research. Efforts have been made to

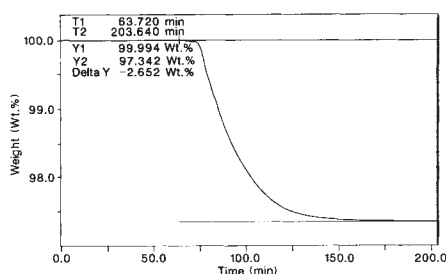


Fig. 1 TGA scan of perovskite obtained using the Perkin-Elmer TGA 7. The sample was heated from 50°C to 200°C at 40°C min⁻¹, held for 60 min, then heated to 1,000°C at a rate of 40°C min⁻¹ and held for 120 min. Courtesy K. Bunshah and R. Budhani.

control the product for better equilibrium weight and diffusibility of oxygen. Changes in oxygen conditioning control have been characterized by various analytical tools, including X-ray diffraction, mass spectrometry and thermal analysis.

Making a superconductor

Making a material which exhibits superconducting properties at relatively high temperatures is surprisingly simple. In fact, secondary school students in the United States and Britain have used published recipes to produce superconductors from materials which are readily available. But to produce the type of materials that have both high critical temperatures and majority phases of superconducting regions, numerous researchers have shown that the oxygen content must be controlled precisely and optimized⁵⁻⁸. For these reasons, elaborate annealing and sintering cycles are currently being used in many research laboratories. These processing cycles typically include annealing and sintering in pure oxygen environments for hours or days at a time, followed by slow cooling in pure oxygen to set temperatures. A final quenching step may be required to achieve the best superconducting properties. The goal of these investigations is to determine the optimum cycles for the production of very pure, high critical temperature superconducting materials for widespread use.

Materials characterization

The processing conditions for synthesizing superconducting materials of high purity and critical temperature are crucial, making necessary laboratory techniques for both the modelling of the process and the measurement of the effects of various different annealing cycles. Thermogravi-

metric analysis (TGA) has proven to be a very useful technique in this area. TGA is a thermal analysis technique that directly measures the change in the mass of a material as the material is heated, cooled or held at a constant temperature in a controlled atmosphere. With a TGA instrument, a researcher is able to control the sample temperature and sample atmosphere, allowing sophisticated modelling of the process on a micro-scale in the TGA instrument.

The characterization of previously prepared materials that exhibit superconducting properties can also be performed using TGA to determine such parameters as the temperature dependence or the equilibrium oxygen content of the sample. Figure 1 demonstrates the use of TGA in the analysis of a perovskite sample that had been quenched previously by exposure to an oxygen atmosphere at 925°C. The experiment involved heating the sample in an argon atmosphere and holding it isothermally at 1,000°C. At this temperature, the sample began to lose weight (with a total weight loss of 2.652 per cent) and finally reached a steady state. The weight loss has been attributed to the loss of oxygen within the matrix. Studies such as these yield important information about the stoichiometry of the oxide states of superconducting materials.

New approaches

Several other thermal analysis techniques are presently available that may be adapted for use in the characterization of superconductors. Differential scanning calorimetry (DSC) instruments operate from -175°C to 730°C, and measure the heat flow into or out of a sample as a function of temperature or time. The specific heat of a superconductor changes when the material is heated or cooled through the temperature where it becomes superconducting. This transition stage is generally known as a lambda transition, because the change in specific heat looks very much like the Greek letter lambda when graphed. Once superconducting materials with transition temperature higher than -175°C — the current lower temperature range of DSC instruments — are developed, DSC will readily find application in the measurement of their specific heats.

Thermomechanical analysis (TMA) operates from -175°C to 1,000°C, and measures the dimensional change in a

sample as a function of temperature or time. Preliminary experiments in our laboratories indicate promise for the use of a magnet coupled to a TMA instrument for the study of the Meissner effect.

Differential thermal analysis (DTA) monitors the heat flux into or out of a sample as a function of temperature, from room temperature to 1,500°C. Initial results in our laboratory on an experimental material have indicated a series of high temperature phase transitions in the region of 800°C to 1,100°C. The study of these phase transitions may lead to a better understanding of the phase behaviour and structural nature of superconducting compounds.

The range of thermal analysis techniques contains many applications that may find future use in the study of superconductors. The use of thermal analysis instrumentation in the characterization of superconductors is expected to increase rapidly as higher temperature materials become available. □

William Brennan, Michael DiVito, Robert Culmo and Colin Williams are at Perkin-Elmer Corporation, 761 Main Avenue, Norwalk, Connecticut 06859-0131 USA. For more information, fill in reader service number 100.

1. Dagni, R. *Chem. Engng News* 65, no. 19, 7-17 (1987).
2. Grant, P. *New Scientist* 7, 36-39 (1987).
3. Strobel, P., Capponi, J.J., Chailout, C., Marezio, M. & Tholence, J.L. *Nature* 327, 306-308 (1987).
4. Dagni, R. *Chem. Engng News* 65, no. 37, 33-35 (1987).
5. Ourmazd, A. *et al. Nature* 327, 308-310 (1987).
6. Morris, D.E. *et al.* 1987 Spring Meeting Materials Research Society, Abstr. No. EA-11.
7. Fisher, G. & Schober, M. *Ceramics Bull.* 66, 1087-1092 (1987).
8. Sawada, H. *et al. Jap. J. appl. Phys.* 26, 11054-11056 (1987).

ADVERTISEMENTS

ULTRA-AMERICA I./E. CO.

Agent for Laboratory Animal Centre, Academy of Military Medical Science, P.R. China, trading high quality animals with quarantine health certificate. Please contact:

**Shop 18
Pelene Arcade
Yue Ko Street
Tin Wan
Hong Kong**

Reader Service No.21

Medwire Div.

stocks a wide variety of

TEFLON COATED*

Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.



Medwire Div.

Affiliate of Sigmund Cohn Corp
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300

*Reg. T.M. of E.I. DuPont

Reader Service No.49

Superconducting neurons?

No, not quite. But this week's feature starts with an overview of products used in the search for room-temperature superconductivity, and ends with several products to be displayed at the Society for Neuroscience meeting in New Orleans, Louisiana, in two weeks.

Johnson Matthey sells a full line of alkaline earth, rare earth, and transition metals and compounds for superconductor research (Reader Service No. 101). Barium and strontium are sold as pure rods, or as carbonate or nitrate compounds, in quantities from 10 g to 2 kg. Rare earths include lanthanum, scandium and yttrium, sold as oxides, powders or ingots ampouled under argon. Copper (II) oxides and nitrates are also available, as are copper powders and rods. Johnson Matthey accommodates requests for customized materials, and ensures confidentiality.

Goodfellow is now selling small random pieces of high-temperature superconducting materials (Reader Service No. 102). The 1-2-3 superconductors are composed of yttrium, barium, copper and oxygen. The pieces are supplied to order, and larger pieces and shapes are under development.

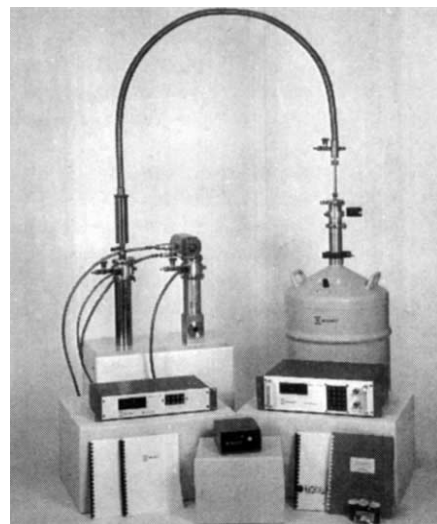
X-ray diffraction may be used to elucidate the crystalline structure of superconducting materials. The model D500 HS high speed X-ray diffraction system from Siemens is designed for measuring the kinetics of a phase transition under the influence of physical variables such as pressure and temperature (Reader Service No. 103). The \$190,000 (US) model D500 HS has a position sensitive proportional counter that allows measurements to be conducted over 50 times faster than with the standard instrument. Siemens' system comes with a process computer, colour graphics, and DIFFRAC 11 software containing programs for diffractometer control, data reduction, indexing, phase determination, quantitative analysis and crystallographic evaluation.

High- T_c tech

For testing high- T_c superconductor samples, Quantum Technology Corp. offers a test system based on its Quantum-cooler closed-cycle cryocooler (Reader Service No. 104). The \$14,200-22,000 (US) system measures the Meissner effect, and electrical resistivity with direct current or alternating current. Measurements may also be made with a magnetic field applied to plot the field dependence of the critical temperature. The electrically operated system does not require liquid helium or nitrogen, and Quantum Technology says it takes only 3 minutes to measure the resistivity/temperature curve of a sample

from 300 K to 10 K. The unit features a dip-tube sample access well filled with helium gas, allowing samples to be inserted and removed without vacuum pumping. Voltage and temperature results may be sent to an X-Y recorder, or to a computer, using an optional interface.

TRI Research has superconductor characterization systems that can be put together to suit a variety of applications (Reader Service No. 105). A complete system includes a cryostat with sample



The components of the TRI Research superconductor characterization system, plus manuals.

holder, electrical and optical access, temperature sensors, temperature read-out and controller, and computer interface, and costs \$12,000-20,000 (US). The cryostat operates on a closed-cycle refrigeration principle, with no need for liquid nitrogen or helium. The temperature range is monitored by a controller over the range of 10-300 K.

A superconductor characterization cryostat is available from APD Cryogenics, Inc. (Reader Service No. 106). The \$20,000-25,000 (US) closed-cycle cryogenic system does not require liquid cryogen, and its sample area can be regulated in temperature from 12 K to over 350 K with ± 0.5 K stability, says the company. The sample tube holds samples as large as 0.5 inches in diameter by 6 inches in length, and an air-lock mechanism permits sample changing without introducing gaseous contaminants, or appreciable warming. The system can be used to perform basic measurements of four-terminal resistance or four-terminal critical current. It can also

be configured using optional accessories for the measurement of Hall voltage, Meissner effect, specific heat, or thermal conductivity.

R.G. Hansen & Associates reports brisk sales of its "university special system" for the characterization of high- T_c superconducting materials (*Reader Service No. 107*). The system is a turn-key laboratory **cryogenic setup**, including a liquid transfer cold head, silicon diode temperature sensor, tip heater, transfer line, temperature controller, vacuum shroud, radiation shield and sample holder — all priced at \$9,960 (US) for academics. When coupled with a researcher's own dewar and liquid cryogens, the system offers a relatively inexpensive option for basic superconductor research.

For under \$20,000 (US), Cryomagnetics, Inc. has a series of **turn-key high- T_c superconductor research systems** (*Reader Service No. 108*). The systems provide the ability to test properties such as critical temperature, magnetic field, and current density, and operate over a 1.8–300 K temperature range, and a 6–12 T magnetic field range.

Thin films

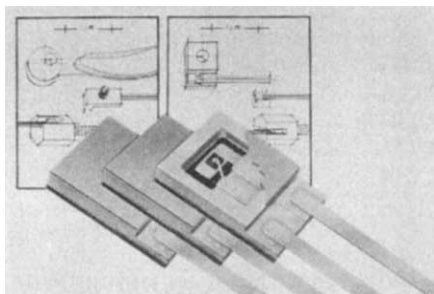
The model 601 **sputtering system** from CVC Products creates thin films from powders by sputtering the target atoms up to the substrate (*Reader Service No. 109*). Sputtering from powders has been shown to be an effective method for thin film preparation, especially when close control of film stoichiometry is required, as in the making of superconducting thin films. CVC Products says the \$125,000 (US) model 601 achieves a film uniformity of ± 5 per cent, and up to four 8 in. circular cathode assemblies can be processed at once. Titanium substrate holders are available that hold fifty-two 50 mm wafers, thirty 75 mm wafers, eighteen 100 mm wafers, seven 125 mm wafers or six 150 mm wafers. Process chamber shields prevent cross-contamination between targets, and are removable for cleaning.

Microscience Inc.'s **Ibex 2000 thin film system** can handle a number of applications: r.f. and d.c. sputtering, magnetron sputtering, superconducting films, ion milling, evaporative coating, ion implantation, reactive ion beam etching and ion surface modification (*Reader Service No. 110*). The instrument has a barrel-style 18- or 24-inch diameter chamber that has a number of large ports to accommodate 10 or 12-inch i.d. modules. Modules are available for each of the applications mentioned above. The system can also be purchased with modifications for *in situ* surface analysis or with load-lock capability. Prices for the Ibex 2000 range from \$100,000 to \$175,000 (US).

The critical temperature

Palm Beach Cryophysics, Inc. has a **thermometer/controller** that operates over a temperature range of 1.5–800 K, with four sensor inputs, in real time (*Reader Service No. 111*). The \$2,595 (US) model 4075 cryogenic thermometer/controller is a microprocessor-controlled instrument with both internal and remote programming. It accepts platinum, rhodium-iron and silicon sensors and has a bar graph error display comprised of 20 segments for ± 100 per cent error about zero.

The model DT-470 series **temperature sensor** from Lake Shore Cryotronics, Inc. was selected as one of the 100 most significant technological developments of 1987 by *Research & Development* magazine (*Reader Service No. 112*). The sensors are



A drawing of Lake Shore's award-winning sensor.

useful in high- T_c superconductor research when used in conjunction with control instrumentation, and have a temperature range of 1.4–475 K. The \$85–450 (US) sensors are made to be interchangeable, and conform to a standard calibration curve to within one of four available accuracy bands. The sensors differ by ± 0.25 K from 2 to 100 K, ± 0.5 K from 100 to 305 K and ± 1.0 K from 305 to 475 K, says Lake Shore.

Magnetic moments

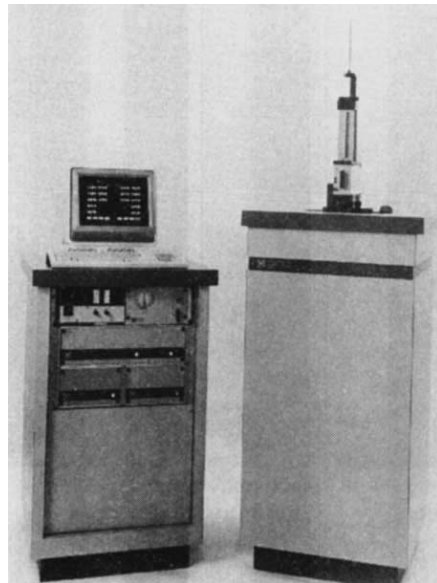
The \$100 **SQUID susceptometer** from Cryogenic Consultants Ltd. performs measurements of the susceptibility and magnetic moment of superconducting materials as a function of temperature and magnetic field (*Reader Service No. 113*). Cryogenic says changes in magnetic moment down to 10^{-8} e.m.u. can be detected over a temperature range of 1.5–300 K. The susceptometer has a wide bore to accommodate large liquid, powder and solid samples. The instrument is magnetically shielded, with a background field of 10^{-7} T with the magnet switched off. According to Cryogenic Consultants, the model S100 SQUID has no magnetic field gradient, eliminating inaccuracies introduced by positioning errors.

A **vibrating sample magnetometer** is available from Oxford Instruments (*Reader Service No. 114*). To use the instrument, the sample under study is

attached to the lower end of a rigid rod and made to oscillate vertically 1.2 mm. If the sample is magnetized, the oscillation will induce an in-phase signal in a set of pickup coils. The instrument is microprocessor-driven, and can be operated at variable temperatures from 4.2 K to 1,000 K.

Janis Research Company, Inc. also has a **turn-key vibrating sample magnetometer** system for the measurement of the magnetic moments of materials as a function of magnetic field and temperature (*Reader Service No. 115*). The system's cryostat covers the temperature range of 1.5–300 K, and magnetic field range of 0–9 T, allowing the transition temperatures for superconducting materials to be determined over the entire range between liquid helium and room temperatures. The complete \$100,000 (US) system includes a model 4500 VSM Princeton Applied Research vibrating sample magnetometer with electronics for sweeping and reversing the field and for automatically controlling and measuring the temperature, plus a variable temperature superconducting magnet cryostat system. All components have an IEEE-488 bus or an RS-232 interface for connection to a computer.

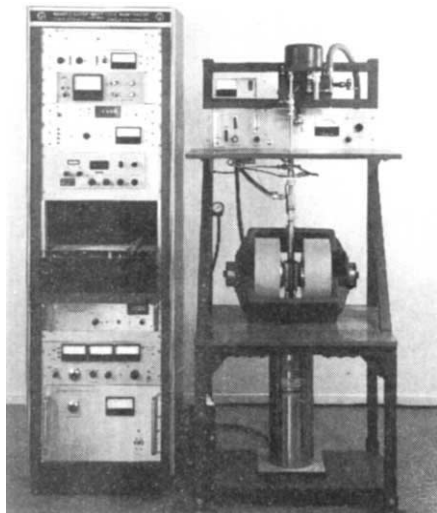
The **Magnetic Property Measurement System** from Quantum Design is capable of resolving magnetic moments as small as 10^{-8} e.m.u. and has a ± 5 T magnetic field, says the company (*Reader Service No. 116*). The \$98,400 instrument employs a SQUID detection system and a superconducting magnet for measuring the magnetic characteristics of materials from 1.8 to 400 K. The menus are operated by touching the screen, and an editor function makes it possible to create files of up to 500 instructions for conducting long series of measurements unattended. Measurements can be tailored for scan-



Computer-driven magnetometer from Quantum Design.

ning, peak-to-peak measurement, and averaging of multiple scans, with the total length, the number of steps and separation between steps selected using the menu.

Magnetometers incorporating Lewis coil sets are being sold by George Associates for the study of high- T_c superconducting materials (*Reader Service No. 117*). Lewis coil sets produce a spatially uniform magnetic field gradient with a magnitude proportional to coil current,



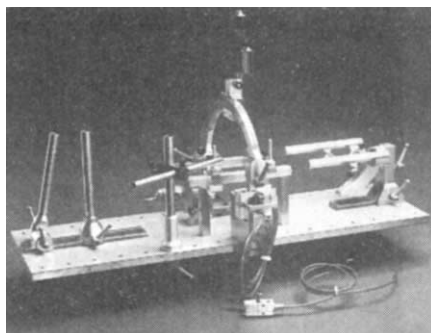
George Associates' magnetometer based on Lewis coils.

and advantages include constant sensitivity and minimized background effects. George Associates says a sample can be loaded and a magnetization curve obtained in under 10 minutes using its series 300 magnetometer systems. The systems have the ability to sweep the magnetic field rapidly — a capability not possible with SQUID magnetometers — to obtain measurements on weakly magnetic samples. The systems have a variable temperature probe cryostat/furnace that operates from 4.2 K to 1,150 K, and uses liquid helium or nitrogen. Different electromagnets provide magnetic fields from 0.75 to 2.5 T. George Associates charges \$97,725 for a complete system.

American Magnetics, Inc. sells a full line of superconducting magnets, cryogenic instruments, and accessories that may be used for the study of high- T_c superconductivity (*Reader Service No. 118*). The company will also perform custom design of **magnet systems** to fulfill special requirements.

For the neuroscientist . . .

A **feline stereotaxic system** for performing *in vivo* electrophysiological experiments is available from Medical Systems Corporation (*Reader Service No. 119*). The \$16,000–28,500 (US) model TRX employs an arc design that the company says facilitates the positioning of instru-

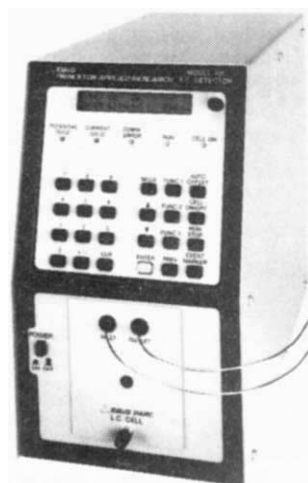


For electrophysiological experiments on feline subjects.

ments along the horizontal coordinates. Precise placement of microelectrodes is aided by one or two stepping motor-driven micromanipulators actuated by an electronic controller. Medical Systems Corporation will be in booths 501–505.

For *in vivo* measurement of oxidizable substances, including ascorbic acid and neurotransmitters such as catecholamines, indolamines and their metabolites, World Precision Instruments offers a **carbon fibre microelectrode** (*Reader Service No. 120*). The \$50 (US) microelectrode has an 8 μm in diameter carbon fibre tip that is 500 μm long. It can be positioned in specific areas of living brain tissue to determine chemical concentrations using voltammetric methods such as normal or differential polarography. The electrode is bonded into a 1.5 mm in diameter and 60 mm long glass tube, and has a bare wire electrical connection. World Precision Instruments is exhibiting a range of electrodes and other instrumentation for electrophysiology in booths 524–528.

The analysis of catecholamines in clinical and biological samples can be accomplished using an **electrochemical detector** for HPLC from EG & G Princeton Applied Research (*Reader Service No. 121*.) The model 400 can be used in



Specializes in the detection of biological samples.

pulse detection mode for carbohydrate analysis, cyclic voltammetry mode for method development, and dual potential mode for the performance of sophisticated detection operations. The detector features an IEEE-488 bus and RS-232 computer interface, plus an auto-offset function that can be triggered by an auto-sampler and a control for pre-electrolysis. The model 400 will be the centrepiece of EG&G's exhibit in booth 313.

Electrical currents

Burleigh has a **three-channel amplifier** for low-voltage piezoelectric actuators for the PZ-300/300M (*Reader Service No. 122*). For \$1,795 (US) the d.c. amplifier offers an output of 0–150 V and 25 mA per channel. It can operate any combination of up to three Burleigh PZL Pushers, PZO Elements or PZS Microstages, individually or simultaneously, says the company. The PZ-300/300M has a fixed gain of 30, automatic overload and short-circuit protection, bandwidth of up to 500 Hz for



Burleigh's amplifier has three channels.

small signals, 0.1 V stability, and maximum noise and ripple of 10 mV, peak to peak. The amplifier complements Burleigh's piezoelectric linear positioning system on display in booth 231.

In booth 321, Bak Electronics, Inc. will be showing its line of biomedical research instrumentation for the neurophysiologist. Among the offerings on display will be the model DIS-1 **amplitude window discriminator** (*Reader Service No. 123*). The \$745 (US) DIS-1 separates single neural spike trains from a multiunit recording containing more than one neural element. The instrument produces a window in both time and amplitude that can be adjusted in height, above or below baseline, and in time up to 5 ms after a trigger point. The trigger point is set by a continuously adjustable level control for either positive or negative slope detection. The signal and both window levels are multiplexed so only one channel of the oscilloscope is necessary, and two DIS-1 instruments may be cascaded for two-point discrimination whenever several spikes have very similar waveforms.

Computer connection

A company called NeuroScience has a system for compiling **topographical brain maps** from EEG data (*Reader Service No.*

124). The Brain Imager, on display in booths 454 and 456, is a self-contained portable system for research or clinical use which allows the quantitation and analysis of EEG and evoked potential (EP) signals. The Brain Imager acquires 28 channels of EEG or EP signals, and has 4 extra channels for the measurement of other physiological data. It creates maps from EEG data to give a spatial view of the distribution of electrical activity over the scalp, with changes in signal voltage and frequency shown as changes in colour. The software provided allows individual and group averages, standard deviations and other statistical averages to be calculated. Stored files can be edited to remove artefact. The \$99,500 (US) Brain Imager includes computer and monitor, printer, optical disk module, impedance tester, somatosensory stimulator, auditory and visual stimulators, self-contained amplifying system, and software.

A software package from Eutectic Electronics turns a properly configured Zeiss microscope and an IBM PC into a tool for **neuron tracing** (Reader Service No. 125). The VDP III Neuron Tree Reconstruction System digitizes microscope data and follows continuous "tree figures" where changes along the Z axis occur within each section. The program traces, edits, displays, plots and stores tree structures, and includes a library of statistical and mathematical morphological summaries. Data can be merged from sequential tissue sections, and 3-dimensional reconstructions of complete, partial or multiple structures can be displayed. Morphometric data such as fibre length and cell volume can be calculated and graphed. Eutectic Electronics sells the software separately, or as part of a complete system, including microscope, computer and plotter, to be exhibited in booths 416-420. A free electronic mail/bulletin board service comes with purchase of an entire system.

ADVERTISEMENT

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures aerobic or anaerobic.

Low Cost



Eliminates evaporation loss and contamination. Easy to use - flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.**

Reader Service No. 174

Biographics, Inc. sells a **computer-aided reconstruction package** (CARP) that can be used to create three-dimensional images of brain sections from photographs, autoradiographs, tomographic scans, and ultrasonography, magnetic resonance imaging or microscope data (Reader Service No. 126). The CARP program stores graphical information such as perimeter descriptions, object positions, and intensities, and viewing parameters such as viewpoints, windows, and eye position in one data file. Quantitative operations include assessment of surface area and volume from solid body models, and three-dimensional distributions of positional information stored in serial sections. The CARP software comes as part of a Neuroanatomy Workstation, including computer, video camera, and image processing and graphics boards, for \$50,000 (US). Biographics is in booth 756.

Probes and reagents

ω -Conotoxin GVIA, a **neurotoxin probe** for neuronal Ca^{2+} channels, is available from Bachem, Inc. (Reader Service No. 127). ω -Conotoxin GVIA is a 27-amino acid peptide, isolated from the venom of *Conus geographicus*, that has been shown to block voltage-activated Ca^{2+} uptake in chick brain synaptosomes. Bachem says data suggest that ω -conotoxin GVIA is a more specific probe for neuronal calcium channels than dihydropyridines. Bachem, exhibiting in booth 430, sells purified synthetic ω -conotoxin for \$200 (US) for 0.5 mg.

Boehringer Mannheim Biochemicals' **1987/1988 general catalogue** hit the streets in July, and the company is sure to have copies on hand in booths 200 and 202 (Reader Service No. 128). The free catalogue lists over 1,800 products for molecular biology, immunology, cell biology and biochemistry, including 300 new products. Each entry contains a detailed description of the product, plus analysis data, storage information and possible applications. The catalogue's expanded table of contents, alphabetical product listings, quick reference tabs and index of related products makes finding needed products faster and easier.

Specimen preparation

The Polaron division of BioRad has come up with a new twist on **specimen preparation** for microscopy-microwave processing. BioRad says the method reduces tissue preparation time by 90 per cent over traditional methods. The model H2500 microwave processor uses microwave energy to stabilize unfixed tissue for both light and electron microscopy (Reader Service No. 129). The processor features a programmable temperature control, pre-set timer and elapsed time LED display for reproduc-

ible results. A probe is supplied for accurate temperature control to $\pm 1^\circ\text{C}$. The processor can also be used to enhance fixation and decrease processing time for paraffin embedment. BioRad's Microscience division will be in booth 402.

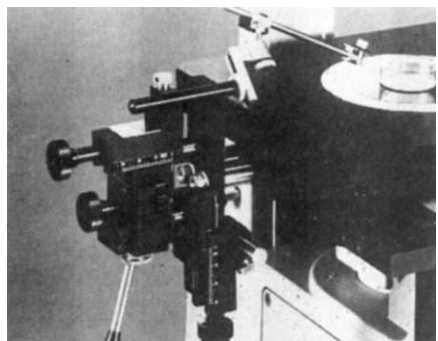
Reichert-Jung has a new **cryosectioning system** that it says offers a significant



A cryosectioning system that's easy on liquid nitrogen.

reduction in liquid nitrogen consumption (Reader Service No. 130). The model FC4E has an ergonomically designed cryochamber with an increased working surface over the previous model FC4D. The unit features a rapid cooling function that allows cryosectioning at temperatures down to -190°C within minutes. Digital displays give the temperature of the specimen, knife and chamber simultaneously. The cryochamber has an open top for unrestricted access to the knife and specimen, and a continuous flow of nitrogen keeps frost from forming. The model FC4E cryosectioning system can be used with the Ultracut E ultramicrotomy system, on display in booths 329-333.

A mechanical **joystick micromanipulator** is new from Narshige USA, in booth 432 (Reader Service No. 131). The \$770 (US) model MN-151 is a lightweight 3-dimensional micromanipulator that fits



Narshige's micromanipulator fits most microscopes.

most microscopes. The model MN-151's joystick allows movements in the X and Y coordinates simultaneously. Independent vernier controls provide movements in X, Y and Z planes over distances of 25 mm.

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

Graduate trends — the latest news

Richard Pearson

The 'shake out' of jobs in the City of London — accentuated by recent events in the stock markets — is one new trend this year.

EACH autumn, this column reviews the latest trends in the graduate labour market as a precursor to the annual Milk Round through which many UK companies conduct their graduate recruitment activities. Traditionally, in the spring term of each year, 500 or more companies advertise vacancies, interview candidates and, they hope, secure their share of the graduates in the summer. Although this remains the dominant source of recruitment, more and more employers are recruiting throughout the year, with more good-quality vacancies appearing at the summer recruitment fairs. In order to attract the best students, many companies are also now starting recruitment in the autumn term, challenging the accountants who have traditionally recruited before Christmas and have thus been able to steal a march on their competitors. The fact that the accountants now take one in nine of university graduates going into employment has led many other employers to follow their lead.

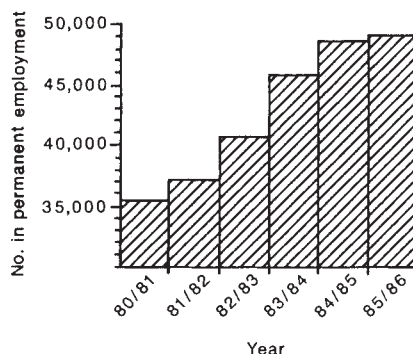
Demand for graduates has continued to rise since 1984, when the economy was pulling out of the recession and the watchword of this column was 'cautious growth in demand for graduates', to be followed in subsequent years by 'a seller's market for graduates' in 1985 and 'competition between employers is becoming intense' in 1986. This year promises to be another record. These ever-improving prospects have been brought about both by an increase in demand and by a reduction in university graduate output over the past three years resulting from the cuts of 1981–82. In 1986 for example, the total output of graduates was about 116,000 at first-degree level of whom 71,000 came from the universities, a figure which had fallen from a peak of 74,000 in 1983, with a compensating rise in output from the polytechnics and other colleges.

There are also some 26,000 postgraduates, of whom more than half are UK residents. But the figures include many postgraduates receiving their degrees when they have already been in employment for several years and exclude a very large but unknown number who fail to complete their doctorates. Another statistical problem has been the change in subject classifications which took place this year which, while helpful for the future, makes it difficult to compare trends in the output of first-degree graduates in different subjects.

Percentages of graduates in permanent employment, by subject

Pharmacy	96.3	88.8
Electrical engineering	82.7	79.3
Mechanical engineering	79.3	81.1
Production engineering	80.4	77.9
Biological sciences	54.0	47.0
Mathematics	82.5	71.0
Chemistry	49.6	47.1
Psychology	50.6	52.1

Source: First Destination Statistics.



First-degree graduates in permanent employment six months after graduation.

From within a broadly static output, the numbers of first-degree graduates going directly into employment has risen for the fifth year running and is set to rise again in 1988 (see figure). Last year also saw a continuation of the fall in the graduate unemployment rate. The precise meaning of these figures is, however, hotly debated. For example, while university graduates are less likely to be unemployed (7.3 per cent in 1986) than their polytechnic counterparts (11.7 per cent in 1986), maintaining a traditional differential, in many subjects polytechnic graduates are more likely to go into permanent employment (see table). This apparent paradox is explained by the greater likelihood that university graduates will go on to further training by way of a postgraduate course in teaching for example, or into research. Another reading of the figures shows only 51.4 per cent of Cambridge graduates getting permanent jobs on graduation against a national average of 60.6 per cent and 77.1 per cent for Brunel.

Demand for graduates in 1988 is expected to be 5 per cent higher than in the year just completed, although this increase is not uniform. Recruitment of new graduates to electronics and computing has still to recover fully from the downturn of last year and the demand by many industrial

employers is continuing to fluctuate. The major growth points are among the newer recruiters, particularly retailers, while solicitors are becoming a major force in the market. The largest recruiters, however, remain accountants, several of whom are seeking 300 or more graduates and one of whom is seeking nearly 1,000.

On the salary front, the start of the shakeout in the City of London may lead to a moderation in the high salaries of £20,000 or more being paid to some graduates, but there are still many City employers and management consultancies paying £13,000 or more. The numbers receiving these salaries are, however, relatively small, totalling a couple of thousand perhaps, and the salary range that most graduates can expect in 1988 is likely to be between £7,000 and £11,000 with an average of just over £9,000. The highest salaries are available mainly in London and are being paid by management and business services companies and those in the oil, chemicals and electronics sectors, while the lower figures are usually associated with jobs involving heavy initial training such as in civil engineering and accountancy, although it is mainly the smaller accountancy practices that offer the lower salaries. Subject differences, where they apply, continue to favour the electronics graduates and those with engineering and computing skills.

Premiums for postgraduate study, where they exist, remain small unless the research is especially relevant, when salaries are set on a one-off basis and may reach £12,000 or more. More usually, the premium, if paid, is around £500 — less than would have been achieved after an equivalent period by a direct entry first-degree graduate. London differentials are averaging around £900 and in some cases reach £2,000, but are not always explicitly identified as many London employers do not offer jobs outside the capital. The principal message on the salary front is one of diversity, with enormous differences between the highs and lows, although the majority of graduates are likely to start on salaries well under £10,000. There are plenty of jobs available and career choice should take account of training and long-term prospects as well as intrinsic interest and salary.

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.